

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 162.447 Seconds
(without alignments)
359.949 Million cell updates/sec

Title: US-10-057-813-14
Perfect score: 861
Sequence: 1 MEHYRKAGSVLPAPSPMPQ.....SMPSSCGPRRRRRARDTRS 163

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	861	100.0	163	4	AAG89339	Aag89339 Human sec
2	861	100.0	163	6	ABU00014	Abu00014 Human nov
3	861	100.0	219	8	ADM90988	Adm90988 Human pha
4	813	94.4	163	2	AAY36104	Aay36104 Extended
5	813	94.4	163	8	ADP19412	Adp19412 Human sec
6	402	46.7	79	2	AAY11950	Aay11950 Human 5'
7	395	45.9	78	3	AAG03599	Aag03599 Human sec
8	339	39.4	211	4	AAM38973	Aam38973 Human pol
9	232.5	27.0	205	4	ABB64410	Abb64410 Drosophil
10	206	23.9	350	3	AAG16838	Aag16838 Arabidops
11	206	23.9	350	3	AAG51875	Aag51875 Arabidops
12	206	23.9	350	3	AAG51847	Aag51847 Arabidops
13	206	23.9	350	8	ADN72883	Adn72883 Thale cre
14	206	23.9	350	8	ADN72183	Adn72183 Thale cre
15	206	23.9	356	3	AAG51846	Aag51846 Arabidops
16	198	23.0	267	3	AAG33136	Aag33136 Zea mays
17	192	22.3	259	3	AAG34072	Aag34072 Zea mays
18	192	22.3	313	3	AAG34071	Aag34071 Zea mays
19	186	21.6	239	3	AAG33137	Aag33137 Zea mays
20	171.5	19.9	314	3	AAG51848	Aag51848 Arabidops
21	171.5	19.9	314	3	AAG16839	Aag16839 Arabidops
22	171.5	19.9	314	3	AAG51876	Aag51876 Arabidops
23	155.5	18.1	207	3	AAG34073	Aag34073 Zea mays
24	155.5	18.1	298	3	AAG51877	Aag51877 Arabidops
25	155.5	18.1	298	3	AAG16840	Aag16840 Arabidops

26	154	17.9	215	3	AAG33138	Aag33138 Zea mays
27	150.5	17.5	99	3	AAG36977	Aag36977 Arabidops
28	135	15.7	74	3	AAG36978	Aag36978 Arabidops
29	104	12.1	124	8	ADP29604	Adp29604 Human sec
30	104	12.1	124	8	ADP29613	Adp29613 Human sec
31	104	12.1	124	8	ADP29615	Adp29615 Human sec
32	104	12.1	124	8	ADP29610	Adp29610 Human sec
33	104	12.1	124	8	ADP29607	Adp29607 Human sec
34	91.5	10.6	791	4	ABG23210	Abg23210 Novel hum
35	91.5	10.6	950	7	ABO77181	Abo77181 Pseudomon
36	91	10.6	345	7	ABO70060	Abo70060 Pseudomon
37	91	10.6	921	7	ABO68758	Abo68758 Pseudomon
38	90.5	10.5	330	7	ABO70617	Abo70617 Pseudomon
39	90	10.5	429	7	ABO71617	Abo71617 Pseudomon
40	89.5	10.4	432	7	ABO74876	Abo74876 Pseudomon
41	88	10.2	393	4	AAU27871	Aau27871 Fruit fly
42	87	10.1	374	7	ABO80084	Abo80084 Pseudomon
43	86.5	10.0	174	6	ADA55419	Ada55419 Human pro
44	86.5	10.0	359	4	AAM25759	Aam25759 Human pro
45	86.5	10.0	371	4	ABG05200	Abg05200 Novel hum

ALIGNMENTS

RESULT 1
AAG89339
ID AAG89339 standard; protein; 163 AA.
XX
AC AAG89339;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 459.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
(GEST) GENSET.
Dumas Milne Edwards J, Bougueleret L, Jobert S;
WPI; 2001-367870/38.
N-PSDB; AAH64942.
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Claim 21; Page 910-911; 921pp; English.
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention
XX
SQ Sequence 163 AA;

Query Match 100.0%; Score 861; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.2e-82;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPD TGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPD TGLDPLTVRRHVPVAVWL 120
QY 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSCGPRSRRRARDTRS 163
Db 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSCGPRSRRRARDTRS 163

RESULT 2
ABU00014
ID ABU00014 standard; protein; 163 AA.
XX
AC ABU00014;

XX
DT 17-JAN-2003 (first entry)
XX
DE Human novel polypeptide #107.

XX Human; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.

XX Homo sapiens.
XX
PN WO200274961-A1.

PD 26-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US005109.

PF 15-MAR-2001; 2001US-00810173.
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2003-040556/03.
DR N-PSDB; ABX05092.
XX

PT New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.

XX Claim 9; SEQ ID NO 633; 235pp; English.
PS
XX The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations

CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABG99888-ABG9989 and ABU00010-ABU00433 represent human
CC polypeptides of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied by the European Patent Office
XX

SQ Sequence 163 AA;
Query Match 100.0%; Score 861; DB 6; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.2e-82;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPD TGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPD TGLDPLTVRRHVPVAVWL 120
QY 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSCGPRSRRRARDTRS 163
Db 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSCGPRSRRRARDTRS 163

RESULT 3
ADM90988
ID ADM90988 standard; protein; 219 AA.
XX
AC ADM90988;

XX
DT 03-JUN-2004 (first entry)
XX
DE Human pharmaceutically useful protein SeqID 381.

XX human; cancer; haematopoiesis; thrombosis; anaemia;
KW cardiovascular disorder; ischaemic heart disease;
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;
KW cystic fibrosis; chronic renal failure; glomerulopathy;
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
KW HIV infection; systemic lupus erythematosus; endocrine system;
KW diabetes mellitus; epilepsy; Alzheimer's disease;
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
KW fungal; parasitic; viral infection; cytostatic; anticoagulant;
KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;
KW immunosuppressive; antiallergic; dermatological; antirheumatic;
KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;
KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;
KW gene therapy; vaccine.

XX Homo sapiens.
XX
PN WO2004020595-A2.
XX
XX 11-MAR-2004.

XX 28-AUG-2003; 2003WO-US027107.
PF
XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410951P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411111P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Hestir K;
XX
DR WPI; 2004-257410/24.
DR N-PSDB; ADM90779, ADM91197.
XX
PT New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 381; 254pp; English.
XX
XX This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC anti-anaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polypeptide is a human protein of the invention.
XX
SQ Sequence 219 AA;
Query Match 100.0%; Score 861; DB 8; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.8e-82;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 57 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 116
QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 117 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 176
QY 121 LSRDPLDPNECGYQPPGAPFGLGMPSSSCGPRRRRRARDTRS 163

Db 177 LSRDPLDPNECGYQPPGAPFGLGMPSSSCGPRRRRRARDTRS 219
RESULT 4
AAAY36104
ID AAY36104 standard; protein; 163 AA.
XX
AC AAY36104;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein sequence, SEQ ID NO. 489.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease.
XX
OS Homo sapiens.
XX
PN WO9931236-A2.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB002122.
XX
PR 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
DR N-PSDB; AAX97788.
XX
PT New isolated human secreted proteins.
XX
PS Claim 9; Page 423-424; 516pp; English.
XX
CC This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases
XX
SQ Sequence 163 AA;
Query Match 94.4%; Score 813; DB 2; Length 163;
Best Local Similarity 94.5%; Pred. No. 1.4e-77;
Matches 154; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLOTEDSWVPXSPDTGLXPLTVRRHVPAXVWL 120
QY 121 LSRDPLDPNECGYQPPGAPFGLGMPSSSCGPRRRRRARDTRS 163

Qy 61 GRAAGKAVSCAEIVKRRVP 79
|||
Db 61 GRAAGKAVSCAEIVKRRVP 79

RESULT 7	
AAG03599	
ID	AAG03599 standard; protein; 78 AA.
XX	
AC	AAG03599;
XX	
DT	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein, SEQ ID NO: 7680.

The present sequence is a polypeptide encoded by one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

Sequence 78 AA;

RESULT 8
AAM38973

AAM38973 standard; protein; 211 AA.
AAM38973;
22-OCT-2001 (first entry)
Human polypeptide SEQ ID NO 2118.
Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
Homo sapiens.
WQ200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US034263.
23-DEC-1999; 99US-00471275.
21-JAN-2000; 2000US-00488725.
25-APR-2000; 2000US-00552317.
20-JUN-2000; 2000US-00598042.
19-JUL-2000; 2000US-00620312.
03-AUG-2000; 2000US-00653450.
14-SEP-2000; 2000US-00662191.
19-OCT-2000; 2000US-00693036.
29-NOV-2000; 2000US-00727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
N-PSDB; AAI58129.
Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
Example 4; SEQ ID NO 2118; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Query Match 39.4%; Score 339; DB 4; Length 211;
Best Local Similarity 41.9%; Pred. No. 2.5e-27;
Matches 72; Conservative 26; Mismatches 56; Indels 18; Gaps 3;

PR	12-AUG-1999;	99US-0148341P;
PR	13-AUG-1999;	99US-0148565P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149368P;
PR	17-AUG-1999;	99US-0149175P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	30-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	23-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	06-OCT-1999;	99US-0157865P;
PR	07-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0159293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159329P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	14-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;
PR	21-OCT-1999;	99US-0160768P;
PR	21-OCT-1999;	99US-0160770P;
PR	21-OCT-1999;	99US-0160814P;
PR	21-OCT-1999;	99US-0160815P;
PR	22-OCT-1999;	99US-0160980P;
PR	22-OCT-1999;	99US-0160981P;
PR	22-OCT-1999;	99US-0160989P;
PR	25-OCT-1999;	99US-0161404P;
PR	25-OCT-1999;	99US-0161405P;
PR	25-OCT-1999;	99US-0161406P;
PR	26-OCT-1999;	99US-0161359P;
PR	26-OCT-1999;	99US-0161360P;
PR	26-OCT-1999;	99US-0161361P;
PR	28-OCT-1999;	99US-0161920P;
PR	28-OCT-1999;	99US-0161992P;
PR	28-OCT-1999;	99US-0161993P;
PR	29-OCT-1999;	99US-0162142P;

Query Match 23.9%; Score 206; DB 3; Length 350;

Query Match 23.5%; Score 206; DB 3;
Best Local Similarity 29.9%; Pred. No. 5.4e-13;

Seq. ID	Seq. Name	Seq. Length	Seq. Type	Seq. Source	Seq. Accession	Seq. Date	Seq. Status	Seq. Notes
1	Seq. 1	100	Protein	Human	NC_000001	2000	Active	Seq. 1
2	Seq. 2	100	Protein	Human	NC_000002	2000	Active	Seq. 2
3	Seq. 3	100	Protein	Human	NC_000003	2000	Active	Seq. 3
4	Seq. 4	100	Protein	Human	NC_000004	2000	Active	Seq. 4
5	Seq. 5	100	Protein	Human	NC_000005	2000	Active	Seq. 5
6	Seq. 6	100	Protein	Human	NC_000006	2000	Active	Seq. 6
7	Seq. 7	100	Protein	Human	NC_000007	2000	Active	Seq. 7
8	Seq. 8	100	Protein	Human	NC_000008	2000	Active	Seq. 8
9	Seq. 9	100	Protein	Human	NC_000009	2000	Active	Seq. 9
10	Seq. 10	100	Protein	Human	NC_000010	2000	Active	Seq. 10
11	Seq. 11	100	Protein	Human	NC_000011	2000	Active	Seq. 11
12	Seq. 12	100	Protein	Human	NC_000012	2000	Active	Seq. 12
13	Seq. 13	100	Protein	Human	NC_000013	2000	Active	Seq. 13
14	Seq. 14	100	Protein	Human	NC_000014	2000	Active	Seq. 14
15	Seq. 15	100	Protein	Human	NC_000015	2000	Active	Seq. 15
16	Seq. 16	100	Protein	Human	NC_000016	2000	Active	Seq. 16
17	Seq. 17	100	Protein	Human	NC_000017	2000	Active	Seq. 17
18	Seq. 18	100	Protein	Human	NC_000018	2000	Active	Seq. 18
19	Seq. 19	100	Protein	Human	NC_000019	2000	Active	Seq. 19
20	Seq. 20	100	Protein	Human	NC_000020	2000	Active	Seq. 20
21	Seq. 21	100	Protein	Human	NC_000021	2000	Active	Seq. 21
22	Seq. 22	100	Protein	Human	NC_000022	2000	Active	Seq. 22
23	Seq. 23	100	Protein	Human	NC_000023	2000	Active	Seq. 23
24	Seq. 24	100	Protein	Human	NC_000024	2000	Active	Seq. 24
25	Seq. 25	100	Protein	Human	NC_000025	2000	Active	Seq. 25
26	Seq. 26	100	Protein	Human	NC_000026	2000	Active	Seq. 26
27	Seq. 27	100	Protein	Human	NC_000027	2000	Active	Seq. 27
28	Seq. 28	100	Protein	Human	NC_000028	2000	Active	Seq. 28
29	Seq. 29	100	Protein	Human	NC_000029	2000	Active	Seq. 29
30	Seq. 30	100	Protein	Human	NC_000030	2000	Active	Seq. 30
31	Seq. 31	100	Protein	Human	NC_000031	2000	Active	Seq. 31
32	Seq. 32	100	Protein	Human	NC_000032	2000	Active	Seq. 32
33	Seq. 33	100	Protein	Human	NC_000033	2000	Active	Seq. 33
34	Seq. 34	100	Protein	Human	NC_000034	2000	Active	Seq. 34
35	Seq. 35	100	Protein	Human	NC_000035	2000	Active	Seq. 35
36	Seq. 36	100	Protein	Human	NC_000036	2000	Active	Seq. 36
37	Seq. 37	100	Protein	Human	NC_000037	2000	Active	Seq. 37
38	Seq. 38	100	Protein	Human	NC_000038	2000	Active	Seq. 38
39	Seq. 39	100	Protein	Human	NC_000039	2000	Active	Seq. 39
40	Seq. 40	100	Protein	Human	NC_000040	2000	Active	Seq. 40
41	Seq. 41	100	Protein	Human	NC_000041	2000	Active	Seq. 41
42	Seq. 42	100	Protein	Human	NC_000042	2000	Active	Seq. 42
43	Seq. 43	100	Protein	Human	NC_000043	2000	Active	Seq. 43
44	Seq. 44	100	Protein	Human	NC_000044			

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60

Db 1 MDKYQR-----VVKPKADTPIDANEIRITSGRARNYITYAMTLLQDKGSTVEVFKAM 533

QY		61	GRAAGKAVSCAEIVRRVPGHQLTKLRFLOTEDSWVPASPDITGLDPLTVRRHVPAVVVL	129
Dd		54	GRAINKTIVTELIKRRIIPDLHQNTSIGSTDITDTWEPT--EEGLPLETRHVSMTIT	111
QY		121	LSDPLDPNEGYQPP-----GAPPGLGMPSSSCGPRSRRAR	159
Dd		112	LSKIELNTSSVGVCPIPIELVKPMGDIDYEGREGSPGGRG-----RGRGRGR	159
 RESULT 12 AAG51847				
ID	AAG51847	standard; protein; 350 AA.		
AC	XX			
CC	XX			
DD	XX			
DT	18-OCT-2000	(first entry)		
DE	XX			
KW	Arabidopsis thaliana	protein fragment SEQ ID NO: 65844.		
KW	Protein identification;	signal transduction pathway; metabolic pathway;		
KW	hybridisation assay;	genetic mapping; gene expression control; promoter;		
KW	termination sequence.			
OS	Arabidopsis thaliana.			
PN	EPI033405-A2.			
PF	06-SEP-2000.			
PX	25-FEB-2000;	2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.		
PR	05-MAR-1999;	99US-0123180P.		
PR	09-MAR-1999;	99US-0123548P.		
PR	23-MAR-1999;	99US-0125788P.		
PR	25-MAR-1999;	99US-0126264P.		
PR	29-MAR-1999;	99US-0126785P.		
PR	01-APR-1999;	99US-0127462P.		
PR	06-APR-1999;	99US-0128234P.		
PR	08-APR-1999;	99US-0128714P.		
PR	16-APR-1999;	99US-0129845P.		
PR	19-APR-1999;	99US-0130077P.		
PR	21-APR-1999;	99US-0130449P.		
PR	23-APR-1999;	99US-0130510P.		
PR	23-APR-1999;	99US-0130891P.		
PR	28-APR-1999;	99US-0131449P.		
PR	30-APR-1999;	99US-0132048P.		
PR	30-APR-1999;	99US-0132407P.		
PR	04-MAY-1999;	99US-0132484P.		
PR	05-MAY-1999;	99US-0132485P.		
PR	06-MAY-1999;	99US-0132486P.		
PR	06-MAY-1999;	99US-0132487P.		
PR	07-MAY-1999;	99US-0132863P.		
PR	11-MAY-1999;	99US-0134256P.		
PR	14-MAY-1999;	99US-0134218P.		
PR	14-MAY-1999;	99US-0134219P.		
PR	14-MAY-1999;	99US-0134221P.		
PR	14-MAY-1999;	99US-0134370P.		
PR	18-MAY-1999;	99US-0134768P.		
PR	19-MAY-1999;	99US-0134941P.		
PR	20-MAY-1999;	99US-0135124P.		
PR	21-MAY-1999;	99US-0135353P.		
PR	24-MAY-1999;	99US-0135629P.		
PR	25-MAY-1999;	99US-0136021P.		
PR	27-MAY-1999;	99US-0136392P.		
PR	28-MAY-1999;	99US-0136782P.		
PR	01-JUN-1999;	99US-0137222P.		
PR	03-JUN-1999;	99US-0137528P.		
PR	04-JUN-1999;	99US-0137502P.		
PR	07-JUN-1999;	99US-0137724P.		
PR	08-JUN-1999;	99US-0138094P.		
PR	10-JUN-1999;	99US-0138540P.		
PR	10-JUN-1999;	99US-0138847P.		

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144881P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.9%; Score 206; DB 3; Length 350;

Best Local Similarity 29.9%; Pred. No. 5.4e-13;

Matches 53; Conservative 23; Mismatches 65; Indels 36; Gaps 4;

Qy 1 MEHYRKAGSVELPAPSPMQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60

Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVVFKAM 53

[illegible]

PR	02-AUG-1999;	99US-0146389P
PR	03-AUG-1999;	99US-0147038P
PR	04-AUG-1999;	99US-0147204P
PR	04-AUG-1999;	99US-0147302P
PR	05-AUG-1999;	99US-0147192P
PR	05-AUG-1999;	99US-0147260P
PR	06-AUG-1999;	99US-0147303P
PR	06-AUG-1999;	99US-0147416P
PR	09-AUG-1999;	99US-0147493P
PR	09-AUG-1999;	99US-0147935P
PR	10-AUG-1999;	99US-0148171P
PR	11-AUG-1999;	99US-0148319P
PR	12-AUG-1999;	99US-0148341P
PR	13-AUG-1999;	99US-0148565P
PR	13-AUG-1999;	99US-0148684P
PR	16-AUG-1999;	99US-0149368P
PR	17-AUG-1999;	99US-0149175P
PR	18-AUG-1999;	99US-0149426P
PR	20-AUG-1999;	99US-0149722P
PR	20-AUG-1999;	99US-0149723P
PR	20-AUG-1999;	99US-0149929P
PR	23-AUG-1999;	99US-0149902P
PR	23-AUG-1999;	99US-0149930P
PR	25-AUG-1999;	99US-0150566P
PR	26-AUG-1999;	99US-0150884P
PR	27-AUG-1999;	99US-0151065P
PR	27-AUG-1999;	99US-0151066P
PR	27-AUG-1999;	99US-0151080P
PR	30-AUG-1999;	99US-0151303P
PR	31-AUG-1999;	99US-0151438P
PR	01-SEP-1999;	99US-0151930P
PR	07-SEP-1999;	99US-0152363P
PR	10-SEP-1999;	99US-0153070P
PR	13-SEP-1999;	99US-0153758P
PR	15-SEP-1999;	99US-0154018P
PR	16-SEP-1999;	99US-0154039P
PR	20-SEP-1999;	99US-0154779P
PR	22-SEP-1999;	99US-0155139P
PR	23-SEP-1999;	99US-0155486P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0156458P
PR	29-SEP-1999;	99US-0156596P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158232P
PR	12-OCT-1999;	99US-0158369P
PR	13-OCT-1999;	99US-0159293P
PR	13-OCT-1999;	99US-0159294P
PR	13-OCT-1999;	99US-0159295P
PR	14-OCT-1999;	99US-0159329P
PR	14-OCT-1999;	99US-0159330P
PR	14-OCT-1999;	99US-0159331P
PR	14-OCT-1999;	99US-0159637P
PR	14-OCT-1999;	99US-0159638P
PR	18-OCT-1999;	99US-0159584P
PR	21-OCT-1999;	99US-0160741P
PR	21-OCT-1999;	99US-0160767P
PR	21-OCT-1999;	99US-0160768P
PR	21-OCT-1999;	99US-0160770P
PR	21-OCT-1999;	99US-0160814P
PR	21-OCT-1999;	99US-0160815P
PR	22-OCT-1999;	99US-0160980P
PR	22-OCT-1999;	99US-0160981P
PR	22-OCT-1999;	99US-0160989P
PR	25-OCT-1999;	99US-0161404P
PR	25-OCT-1999;	99US-0161405P
PR	25-OCT-1999;	99US-0161406P
PR	26-OCT-1999;	99US-0161359P
PR	26-OCT-1999;	99US-0161360P
PR	26-OCT-1999;	99US-0161361P
PR	28-OCT-1999;	99US-0161920P

[illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 10:10:14 ; Search time 390 Seconds
(without alignments)
6622.347 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgctgta 492

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_23Sep04:*
1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	492	100.0	660	12	ADM91197	Adm91197 Human DNA
2	492	100.0	660	12	ADM90779	Adm90779 Human DNA
3	492	100.0	894	5	AAH64942	Aah64942 Human sec
4	492	100.0	1089	8	ABX05092	Abx05092 Human nov
5	484	98.4	884	2	AAX97788	Aax97788 Extended
6	484	98.4	884	12	ADP19055	Adp19055 Human sec
7	385.6	78.4	1036	3	AAC95569	Aac95569 Human sec
8	385.6	78.4	1036	8	ADA98058	Ada98058 Human sec
9	385.6	78.4	1036	8	ADA43944	Ada43944 Human sec
10	385.6	78.4	1036	10	ADC20214	Adc20214 Human sec
11	385.6	78.4	1036	10	ADF10630	Adf10630 Human sec
12	373.8	76.0	469	9	ACH33887	Ach33887 Human end
13	368.4	74.9	1168	9	ADA11077	Ada11077 Human cdn
14	368.4	74.9	1168	10	ADJ56254	Adj56254 Human cdn
15	236	48.0	303	3	AAC03605	Aac03605 Human sec
16	236	48.0	305	2	AAX40672	Aax40672 Human sec
17	223.8	45.5	560	10	ABT41218	Abt41218 Toxicity
18	78	15.9	1424	4	AAI59915	Aai59915 Human pol
19	74.6	15.2	1941	4	AAI58129	Aai58129 Human pol
20	74.6	15.2	1941	5	ADQ98335	Adq98335 DNA encod
21	74.6	15.2	1941	9	ADB48095	Adb48095 Novel hum

22	74.6	15.2	2375	12	ADL12363	Adl12363 Human ste
23	68.2	13.9	412	8	ABX52110	Abx52110 Bovine ES
24	44.4	9.0	648	4	ABL15187	Abl15187 Drosophl
25	44.4	9.0	2000	8	ADA71938	Ada71938 Rice gene
c	41	8.3	2000	8	ADA71938	Ada71938 Rice gene
27	40.4	8.2	929	3	AAF15097	Aaf15097 Trichoder
28	39.6	8.0	2816	4	ABL15186	Ab115186 Drosophil
c	39.4	8.0	32189	4	AAL36926	Aal36926 Human mus
c	39.4	8.0	32189	5	ABA16341	Abal6341 Human ner
c	39.4	8.0	32189	5	ABA17492	Abal7492 Human ner
c	39.4	8.0	32189	5	ABA17790	Abal7790 Human ner
c	39.4	8.0	32189	8	ABX59914	Abx59914 cDNA enco
c	39.4	8.0	32189	12	ADJ30664	Adj30664 Human mus
35	37.2	7.6	1338	3	AAC43679	Aac43679 Zea mays
c	37.2	7.6	46819	9	ACD19059	Acd19059 E. coli 0
c	37.2	7.6	46897	10	ADC00585	Adc00585 Enterohae
38	36.6	7.4	3633	12	ADM66939	Adm66939 Murine ad
39	36.2	7.4	500	10	ADE81670	Ade81670 Arabidops
40	36.2	7.4	1053	12	ADN72182	Adn72182 Thale cre
41	36.2	7.4	1053	12	ADN72882	Adn72882 Thale cre
42	36.2	7.4	1071	3	AAC50707	Aac50707 Arabidops
43	36.2	7.4	1603	3	AAC50718	Aac50718 Arabidops
44	36.2	7.4	1605	3	AAC37507	Aac37507 Arabidops
c	36	7.3	11100	10	ADI23891	Adi23891 Streptomy

ALIGNMENTS

RESULT 1
ADM91197
ID ADM91197 standard; DNA; 660 BP.
XX
AC ADM91197;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human DNA encoding a pharmaceutically useful protein SeqID 590.
XX
KW gene; ds; human; cancer; haematopoiesis; thrombosis; anaemia;
KW cardiovascular disorder; ischaemic heart disease;
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;
KW cystic fibrosis; chronic renal failure; glomerulopathy;
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
KW HIV infection; systemic lupus erythematosus; endocrine system;
KW diabetes mellitus; epilepsy; Alzheimer's disease;
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
KW fungal; parasitic; viral infection; cytostatic; anticoagulant;
KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;
KW immunosuppressive; antiallergic; dermatological; antirheumatic;
KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;
KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;
KW gene therapy; vaccine.
OS Homo sapiens.
PN WO2004020595-A2.
XX
PD 11-MAR-2004.
XX
PF 28-AUG-2003; 2003WO-US027107.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410951P.
PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411111P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Hestir K;
XX
DR WPI; 2004-257410/24.
DR P-PSDB; ADM90988.
XX
PT New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 590; 254pp; English.
XX
CC This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC antianaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polynucleotide is a human DNA sequence encoding a
CC polypeptide of the invention.
XX
SQ Sequence 660 BP; 109 A; 229 C; 206 G; 116 T; 0 U; 0 Other;

Query Match 100.0%; Score 492; DB 12; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.9e-120;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db |||||||
QY 61 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db |||||||
QY 229 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288
Db |||||||
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db |||||||
QY 289 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 348
Db |||||||

QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTGCGTGAGATTGTCAAGCGGGGTCCCAGGC 240
Db |||||||
QY 349 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTGCGTGAGATTGTCAAGCGGGGTCCCAGGC 408
Db |||||||
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db |||||||
QY 409 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 468
QY 301 CCTGACACAGGGCTAGACCCCTCAGAGTCGCGCCGCTGTCCTGTCAGTGTGGGTGCTG 360
Db |||||||
QY 469 CCTGACACAGGGCTAGACCCCTCAGAGTCGCGCCGCTGTCCTGTCAGTGTGGGTGCTG 528
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTCACAAACCCCGAGGAGCACCCCT 420
Db |||||||
QY 529 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTCACAAACCCCGAGGAGCACCCCT 588
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGGTCGAGAC 480
Db |||||||
QY 481 ACCCGATCGTGA 492
Db |||||||
QY 649 ACCCGATCGTGA 660

RESULT 2
ADM90779
ID ADM90779 standard; DNA; 660 BP.
XX
AC ADM90779;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human DNA encoding a pharmaceutically useful protein SeqID 172.
XX
KW gene; ds; human; cancer; haematopoiesis; thrombosis; anaemia;
KW cardiovascular disorder; ischaemic heart disease;
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;
KW cystic fibrosis; chronic renal failure; glomerulopathy;
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
KW HIV infection; systemic lupus erythematosus; endocrine system;
KW diabetes mellitus; epilepsy; Alzheimer's disease;
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
KW fungal; parasitic; viral infection; cytostatic; anticoagulant;
KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;
KW immunosuppressive; antiallergic; dermatological; antirheumatic;
KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;
KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;
KW gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2004020595-A2.
XX
PD 11-MAR-2004.
XX
PF 28-AUG-2003; 2003WO-US027107.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410951P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411111P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Hestir K;
XX
DR WPI; 2004-257410/24.
DR P-PSDB; ADM90988.
XX
PT New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 172; 254pp; English.
XX
CC This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC antianaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polynucleotide is a human DNA sequence encoding a
CC polypeptide of the invention.
XX
SQ Sequence 660 BP; 109 A; 229 C; 206 G; 116 T; 0 U; 0 Other;
Query Match 100.0%; Score 492; DB 12; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.9e-120;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db |||||
169 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 228
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db |||||
229 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288
QY 121 GGGTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCCGCATGTAGTGTCTCAGGTTCT 180
Db |||||
289 GGGTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCCGCATGTAGTGTCTCAGGTTCT 348
QY 181 GGCAGGGCTGCAGGAAGGTTGTACAGTGGCTGAGATTGTCAAGCGGCGGGTCCCAGGC 240
Db |||||

Db 349 GGCAGGGCTGCAGGAAAGGCTGTTCAGCTGGCTGAGATTGTCAAGCGGGGGTCCCAGGC 408
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCAGCCTCA 300
Db |||||
409 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCAGCCTCA 468
QY 301 CCTGACACAGGCTAGACCCCTCCTCACAGTGGCGCCCATGTGCTGAGTGTGGGTGCTG 360
Db CCTGACACAGGCTAGACCCCTCCTCACAGTGGCGCCCATGTGCTGAGTGTGGGTGCTG 528
QY 361 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 588
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 648
QY 481 ACCCGATCGTGA 492
Db |||||
649 ACCCGATCGTGA 660
RESULT 3
AAH64942
ID AAH64942 standard; cDNA; 894 BP.
AC AAH64942;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 218.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI WPI; 2001-367870/38.
XX P-PSDB; AAG89339.
DR Full length GENSET human nucleic acids encoding potentially secreted
DR proteins, useful in gene therapy and vaccination against a variety of
XX diseases, and for diagnosis of those diseases.
PS Claim 7; Page 772-773; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which

CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX

SQ Sequence 894 BP; 189 A; 258 C; 255 G; 192 T; 0 U; 0 Other;

Query Match 100.0%; Score 492; DB 5; Length 894;
Best Local Similarity 100.0%; Pred. No. 6.3e-120;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 60
Db |||||||
QY 78 ATGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 137
Db |||||||
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db |||||||
QY 138 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 197
QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
Db |||||||
QY 198 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGTGGCATGTAGTGTCTCAGGTTCT 257
QY 181 GGCAGGGCTGCAGGAAGCTGTGAGCTGCGTGGATGTCAAGCGGGTCCCAGGC 240
Db |||||||
QY 258 GGCAGGGCTGCAGGAAGCTGTGAGCTGCGTGGATGTCAAGCGGGTCCCAGGC 317
QY 241 CTGACCACTGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db |||||||
QY 318 CTGACCACTGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 377
QY 301 CCTGACACAGGGCTAGACCCCTCAGTGGCGGCCATGTGCCTGAGTGTGGGTGCTG 360
Db |||||||
QY 378 CCTGACACAGGGCTAGACCCCTCAGTGGCGGCCATGTGCCTGAGTGTGGGTGCTG 437
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db |||||||
QY 438 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 497
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db |||||||
QY 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCTCGTTCCCGAAGAGGGCTCGAGAC 557
QY 481 ACCCGATCGTGA 492
Db |||||||
QY 558 ACCCGATCGTGA 569

RESULT 4
ABX05092
ID ABX05092 standard; cDNA; 1089 BP.

XX

AC ABX05092;

XX 17-JAN-2003 (first entry)

DT Human novel polynucleotide #107.

XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.

OS Homo sapiens.

XX WO200274961-A1.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-US005109.

XX
PR 15-MAR-2001; 2001US-00810173.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-040556/03.
DR P-PSDB; ABU00014.
XX
PT New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
PS Claim 1; SEQ ID NO 107; 235pp; English.
XX
CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC periodontal disease, osteoarthritis, bone degenerative disorders,
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office
XX
SQ Sequence 1089 BP; 221 A; 306 C; 331 G; 231 T; 0 U; 0 Other;

Query Match 100.0%; Score 492; DB 8; Length 1089;
Best Local Similarity 100.0%; Pred. No. 6.6e-120;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 60
Db |||||||
QY 275 ATGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 334
Db |||||||
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db |||||||
QY 335 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
Db |||||||
QY 395 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGTGGCATGTAGTGTCTCAGGTTCT 454
QY 181 GGCAGGGCTGCAGGAAGCTGTGAGTGTGCGTGGTCAAGCGGGGTCCCAGGC 240
Db |||||||
QY 455 GGCAGGGCTGCAGGAAGCTGTGAGTGTGCGTGGTCAAGCGGGGTCCCAGGC 514
QY 241 CTGACCACTGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db |||||||
QY 515 CTGACCACTGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 574
QY 301 CCTGACACAGGGCTAGACCCCTCAGTGGCGGCCATGTGCCTGAGTGTGGGTGCTG 360
Db |||||||
QY 575 CCTGACACAGGGCTAGACCCCTCAGTGGCGGCCATGTGCCTGAGTGTGGGTGCTG 634
QY 361 CTGACCGGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db |||||||
QY 635 CTGACCGGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 694
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCTCGTTCCCCGAAGAGGGCTCGAGAC 480
Db |||||||
QY 695 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCTCGTTCCCCGAAGAGGGCTCGAGAC 754

QY 481 ACCCGATCGTGA 492
| | | | |
Db 755 ACCCGATCGTGA 766

RESULT 5
AAX97788
ID AAX97788 standard; DNA; 884 BP.
XX
AC AAX97788;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein coding sequence, SEQ ID NO. 353.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
genetic disease; ss.
XX
OS Homo sapiens.
XX
PN WO9931236-A2.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB002122.
XX
PR 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
P-PSDB; AAY36104.
XX
PT New isolated human secreted proteins.
XX
PS Claim 1; Page 423-424; 516pp; English.
XX
CC This sequence represents an extended human secreted protein coding
sequence of the invention. The secreted proteins can be used in treating
or controlling a variety of human conditions. The secreted proteins may
act as cytokines or may affect cellular proliferation or differentiation
or may act as immune system regulators, haematopoiesis regulators, tissue
growth regulators, regulators of reproductive hormones or cell movement
or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
tumour inhibition activity. The DNAs can be used in forensic procedures
to identify individuals or in diagnostic procedures to identify
individuals having genetic diseases resulting from abnormal expression of
the genes corresponding to the extended cDNAs. They are also useful for
constructing a high resolution map of the human chromosomes. They can
also be used for gene therapy to control or treat genetic diseases
XX
SQ Sequence 884 BP; 187 A; 254 C; 238 G; 190 T; 0 U; 15 Other;

Query Match 98.4%; Score 484; DB 2; Length 884;
Best Local Similarity 97.8%; Pred. No. 8.1e-118;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 60
| | | | |
Db 69 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 128
| | | | |
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
| | | | |
Db 129 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 188
| | | | |

QY 121 GGGTTGGCTCTGGGTGGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTTCTCAGGTTCT 180
| | | | |
Db 189 GGGTTGGCTCTGGGTGGGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTTCTCAGGTTCT 248
| | | | |

QY 181 GGCAGGGCTGCAGGAAAGGCTGTCACTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
| | | | |
Db 249 GGCAGGGCTGCAGGAAAGGCTGTCACTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC 308
| | | | |

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
| | | | |
Db 309 CTGCACCAAGCTACCAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 368
| | | | |

QY 301 CCTGACACAGGGCTAGACCCCTCACAGTCCGCCCATGTGCTGCGAGTGTGGGTGCTG 360
| | | | |
Db 369 CCTGACACAGGGCTARACCCCTCACAGTCCGCCCATGTGCTGCGAGTGTGGGTGCTG 428
| | | | |

QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 420
| | | | |
Db 429 CTCASCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 488
| | | | |

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGCGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 480
| | | | |
Db 489 GGCCTGGGTTCCATGCCAGCTCCAGCTGCGCCCTCGTTCCCGAAGAAGGGCTCRARAC 548
| | | | |

QY 481 ACCCGATCGTGA 492
| | | | |
Db 549 ACCCGATCGTGA 560
| | | | |

RESULT 6
ADP19055
ID ADP19055 standard; cDNA; 884 BP.
XX
AC ADP19055;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human secreted polynucleotide #311.
XX
KW Human; secreted protein; gene; ss; genetic disease.
XX
OS Homo sapiens.
XX
PN US2004110939-A1.
XX
PD 10-JUN-2004.
XX
PF 15-OCT-2001; 2001US-00978360.
XX
PR 17-DEC-1998; 98WO-IB002122.
PR 09-FEB-1999; 99WO-IB000282.
PR 21-JUN-2000; 2000WO-IB000951.
PR 15-SEP-2000; 2000US-00663600.
XX
PA (GEST) GENSET SA.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;
PI Duclert A;
XX
DR WPI; 2004-440404/41.
DR P-PSDB; ADP19460.
XX
PT New isolated polynucleotide encoding secreted polypeptide, useful for
PT gene therapy, or in diagnostic procedures to identify individuals having
PT genetic diseases resulting from abnormal expression of the genes.
XX
PS Claim 1; SEQ ID NO 311; 113pp; English.
XX
CC The invention relates to human cDNA sequences that encode human secreted
proteins. The invention also relates to an antibody that specifically
binds to a polypeptide of the invention and a method of binding the
polypeptide to an antibody. The polynucleotides are useful for expressing
the entire secreted proteins which they encode and for distinguishing

CC human tissues and cells from non-human tissues and cells, and for
CC distinguishing between human tissues and cells that do or do not express
CC the polynucleotides comprising the cDNAs. The polynucleotides and
CC polypeptides are useful in forensic procedures or diagnostic procedures
CC to identify individuals with genetic diseases resulting from abnormal
CC expression of the genes corresponding to the cDNAs. The sequences are
CC also useful in gene therapy to control or treat genetic diseases. This
CC sequence represents a human secreted polynucleotide of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.

SQ Sequence 884 BP; 187 A; 254 C; 238 G; 190 T; 0 U; 15 Other;

Query Match 98.4%; Score 484; DB 12; Length 884;
Best Local Similarity 97.8%; Pred. No. 8.1e-118;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
69 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 128
QY 61 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
129 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 188
QY 121 GGGTTGGCTCTGGTGGTGGAGGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
189 GGGTTGGCTCTGGTGGTGGAGGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 248
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
249 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 308
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCCAGACTGAGGACAGTGGTGGTGGTGGT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
309 CTGCACCAAGCTACCAAGCTACGTTTCCTTCCAACTGAGGACAGTGGTGGTGGTGGT 368
QY 301 CCTGACACAGGGTAGACCCCTCAGAGTGGCGCGCCCATGTGCCTGAGTGGTGGTGGT 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
369 CCTGACACAGGGTAGACCCCTCAGAGTGGCGCGCCCATGTGCCTGCAKTGGTGGTGGT 428
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
429 CTCASCCGGGACCCCTGGACCCCAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 488
QY 421 GGCCTGGTTCATGCCAGCTCCAGCTGCGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db GGCCTGGTTCATGCCAGCTCCAGCTGCGGCCCTCGTTCCCGAAGAGGGCTCGARAC 548
QY 481 ACCCGATCGTGA 492
Db ||||||||||||
549 ACCCGATCGTGA 560

RESULT 7
AAC95569
ID AAC95569 standard; cDNA; 1036 BP.

XX AAC95569;

AC AAC95569;

XX 21-FEB-2001 (first entry)

DE Human secreted protein gene 49 SEQ ID NO:59.

XX Human; secreted protein; cytostatic; immunosuppressive; neurotropic;
KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;
KW antifungal; antiparasitic; cardiant; cancer; immune disease; allergy;
KW cardiovascular disorder; wound healing; infection; neurological disease;

XX ss.

XX Homo sapiens.

OS

XX WO200061596-A1.
PN 19-OCT-2000.
XX 06-APR-2000; 2000WO-US008983.
XX 09-APR-1999; 99US-0128703P.
PR 14-JAN-2000; 2000US-0176068P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX Ruben SM, Komatsoulis G;
PI WPI; 2000-611865/58.
XX P-PSDB; AAB52060.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
PT the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases.

XX Claim 1; Page 446; 505pp; English.

XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
CC AAB52103 represent alternative polypeptides encoded by the genes, and
CC amino acid sequences with which they share homology. The genes and
CC proteins have activities dependent on the tissues and cells in which they
CC are expressed. Examples of their activities include cytostatic;
CC immunosuppressive; neurotropic; neuroprotective; antiviral; antiallergic;
CC hepatotropic; antidiabetic; antiinflammatory; antiulcer; vulnery;
CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
CC The secreted proteins, polypeptides, antagonists and agonists may be
CC useful in treating, preventing and/or diagnosing diseases and disorders
CC such as cancer, particularly breast and ovarian cancer, and other cancers
CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases such as viral,
CC bacterial, fungal and parasitic infections may also be treated using the
CC proteins and polynucleotides of the invention. Sequences AAC95512 -
CC AAC95520 and AAB52011 are used in the isolation and characterisation of
CC the proteins and polynucleotides of the invention

XX Sequence 1036 BP; 368 A; 238 C; 235 G; 192 T; 0 U; 3 Other;

Query Match 78.4%; Score 385.6; DB 3; Length 1036;
Best Local Similarity 87.8%; Pred. No. 7.9e-92;
Matches 432; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
44 ATGGAGCACTACCGGAAAGCTGGCTCTGTGGAACCTCCAGACACCTTCTCCGATGCCGCGG 103
QY 61 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
104 CTACCTCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTACTG 163
QY 121 GGGTTGGCTCTGGTGGTGGAGGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
164 GGGTTGGCACTGGTGGTGGATTTGA-GGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 222
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
223 GGTGGGGCTGCAGGGAAGGCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 282
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCCAGACTGAGGACAGTGGTGGTGGTGGT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
283 CTGCACCAAGCTTACCAAGCTGGCTTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGT 342

DR WPI; 2003-184045/18.
XX P-PSDB; ADA44134.
PT A human secreted protein and nucleic acids useful for preparing a
PT diagnostic or pharmaceutical composition for diagnosing or treating
PT diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,
PT retinopathy, neuropathy.
XX
PS Claim 7; SEQ ID NO 132; 701pp; English.
XX
CC The invention relates to novel genes and their fragments which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids and proteins
CC are useful in the diagnosis, treatment and prevention of conditions
CC related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,
CC polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,
CC infection, cataract, renal disorders, or endocrine disorders. The present
CC sequence was used to illustrate the invention.
XX
SQ Sequence 1036 BP; 368 A; 238 C; 235 G; 192 T; 0 U; 3 Other;

Query Match 78.4%; Score 385.6; DB 8; Length 1036;
Best Local Similarity 87.8%; Pred. No. 7.9e-92;
Matches 432; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 44 ATGGAGCACTACCGGAAAGCTGGCTCTGTGGAACCTCCAGCACCTTCTCCGATGCCGCG 103
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCCGAACCTGCTG 120
Db 104 CTACCTCCTGATACCCCTGGAGATGCGGGTCCGAGATGGCAGCAAAATCCGCAACCTACTG 163
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTGCGGCATGTAGTGTCTCAGGTTCT 180
Db 164 GGGCTGGCACTGGGTGATTGGA-GGTGGTGGTGCAGGCGCATGTGGTGTCTCAGGTTCT 222
QY 181 GGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGCGGTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db 223 GGTGCGGCTGCAGGGAAGGCACTGAGTGTGCTGAGATTGTCAAGCGGCGGTACCGGGC 282
QY 241 CTGCACCACTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 300
Db 283 CTGCACCACTTACCAAGCTGCGCTTCTTCAGACCGAGGACAGCTGGGTGCCAGTCTCA 342
QY 301 CCTGACACAGGGCTAGACCCCTCAGAGTGCAGTGCAGGCGCCCATGTGCTGAGTGGTGTG 360
Db 343 CCTGACACAGGCTGATCCCTCAGGTACGCGGCCATGTACCTGCACTGGGTACTG 402
QY 361 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 403 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGCTACAGCCTCCAGGACACCCCT 462
QY 421 GGCCTGGGTTCCATGCCCACTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 463 GGCCTGGGCCCCACATCAAGCTCCAGCTGTGGTCCACGACCCCGAAGAGGGTTCCGAGAC 522
QY 481 ACCCGATCGTGA 492
Db 523 ACCTGGTCTCTAA 534

RESULT 10
ADC20214
ID ADC20214 standard; DNA; 1036 BP.
XX
AC ADC20214;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human secreted protein coding sequence #153.
XX

KW gene therapy; human; secreted protein; haemopoietic disorder;
KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW leukaemia; wound healing; epithelial cell proliferation disorder;
KW immune disorder; autoimmune disorder; asthmatic disorder;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX
XX Homo sapiens.
OS
XX WO200292787-A2.
PN
XX 21-NOV-2002.
PD
XX 26-MAR-2002; 2002WO-US009257.
XX
PF 27-MAR-2001; 2001US-0278650P.
XX 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-129287/12.
DR
XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating hematopoietic or hematologic disorders, e.g.
PT anemia or hemophilia.
PT
XX
PS Claim 1; SEQ ID NO 163; 1512pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful for detecting, preventing, diagnosing, prognosticating, treating
CC or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC wound healing and disorders of epithelial cell proliferation; immune
CC disorders (e.g. autoimmune disorders and asthmatic disorders);
CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and
CC gastroenteritis). The present DNA sequence encodes a human secreted
CC protein of the invention.
XX
SQ Sequence 1036 BP; 368 A; 238 C; 235 G; 192 T; 0 U; 3 Other;

Query Match 78.4%; Score 385.6; DB 10; Length 1036;
Best Local Similarity 87.8%; Pred. No. 7.9e-92;
Matches 432; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 44 ATGGAGCACTACCGGAAAGCTGGCTCTGTGGAACCTCCAGCACCTTCTCCGATGCCGCG 103
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCCGAACCTGCTG 120
Db 104 CTACCTCCTGATACCCCTGGAGATGCGGGTCCGAGATGGCAGCAAAATCCGCAACCTACTG 163
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTGCGGCATGTAGTGTCTCAGGTTCT 180
Db 164 GGGCTGGCACTGGGTGATTGGA-GGTGGTGGTGCAGGCGCATGTGGTGTCTCAGGTTCT 222
QY 181 GGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGCGGTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db 223 GGTGCGGCTGCAGGGAAGGCACTGAGTGTGCTGAGATTGTCAAGCGGCGGTGTACCGGGC 282
QY 241 CTGCACCACTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGGCTCA 300
Db 283 CTGCACCACTTACCAAGCTGCGCTTCTTCAGACCGGAGGACAGCTGGGTGCCAGTCTCA 342

XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene
XX PT mapping, in the recombinant production of protein, or in generating
XX PT antisense DNA or RNA.
XX
XX PS Claim 1; SEQ ID NO 21099; 44pp; English.
XX
XX CC The invention relates to an isolated polynucleotide comprising any one of
XX CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX CC determined by the technique of SBH (sequencing by hybridisation). Also
XX CC included is a purified polypeptide comprising a sequence corresponding to
XX CC a reading frame of the novel polynucleotide. The nucleic acid sequences
XX CC are useful in diagnostics as expressed sequence tags (EST) for
XX CC identifying expressed genes or for physical mapping of the human genome,
XX CC in forensics, in assessing biodiversity, or in identifying mutations
XX CC responsible for genetic disorders and other traits. The nucleotide
XX CC sequences are also useful as hybridisation probes, as oligomers for PCR,
XX CC for chromosome and gene mapping, in the recombinant production of
XX CC protein, or in generating antisense DNA or RNA. The purified polypeptide
XX CC is useful for generating antibodies specific for it. The present sequence
XX CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX SQ Sequence 469 BP; 93 A; 141 C; 149 G; 86 T; 0 U; 0 Other;

Query Match 76.0%; Score 373.8; DB 9; Length 469;
Best Local Similarity 99.2%; Pred. No. 8.5e-89;
Matches 386; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCAGCGCCTTCCCAATGCCCCAG 60
Db |||||
QY 61 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATCGCAACCTGCTG 120
Db |||||
QY 141 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATCGCAACCTGCTG 200
Db |||||
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAG-GTTC 179
Db |||||
QY 180 TGGCAGGGCTGCAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGG 239
Db |||||
QY 261 TGGCAGGGCTGCAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGG 320
Db |||||
QY 240 CCTGCACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTC 299
Db |||||
QY 321 CCTGCACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTC 380
Db |||||
QY 300 ACCTGACACAGGGCTAGACCCCTCAGTGGCGCCGCTGTCCTGAGTGTGGGTGCT 359
Db |||||
QY 381 ACCTGACACAGGGCTAGACCCCTCAGTGGCGCCGCTGTCCTGAGTGTGGGTGCT 440
Db |||||
QY 360 GCTCAGCCGGACCCCTGGACCCCAATG 388
Db |||||
QY 441 GCTCATCCGGGACCCCTGGACCCCAATG 469

RESULT 13
ADA11077/c
ID ADA11077 standard; cDNA; 1168 BP.
XX AC ADA11077;
XX DT 06-NOV-2003 (first entry)
XX DE Human cDNA differentially expressed in colon cancer #1.
XX KW ss; differential expression; colon cancer; cancer; human.
XX OS Homo sapiens.
XX PN US2002160382-A1.
XX PD 31-OCT-2002.
XX PF 11-OCT-2001; 2001US-00981353.
XX PR 11-OCT-2000; 2000US-0239841P.
XX PA (LASE/) LASEK A W.
XX PA (JONE/) JONES D A.
XX PI Lasek AW, Jones DA;
XX WPI; 2003-265756/26.
XX
XX PT New combination comprising cDNAs that are differentially expressed in
XX PT colon disorder, useful for diagnosing, treating, staging or monitoring
XX PT treatment for colon cancers.
XX
XX PS Claim 1; SEQ ID NO 1; 231pp; English.
XX
XX CC The invention relates to a combination comprising cDNAs that are
XX CC differentially expressed in colon disorder. The methods and compositions
XX CC of the present invention are useful for diagnosing, treating, staging or
XX CC monitoring treatment for colon cancer. They are also useful in high
XX CC throughput methods for using cDNAs to detect differential expression of
XX CC nucleic acids in a sample, screening molecules or compounds to identify a
XX CC ligand which specifically binds a cDNA and using a protein to screen
XX CC molecules or compounds to identify at least one ligand which specifically
XX CC binds the protein. The present sequence represents a human cDNA
XX CC differentially expressed in colon cancer.
XX
XX SQ Sequence 1168 BP; 249 A; 346 C; 339 G; 234 T; 0 U; 0 Other;

Query Match 74.9%; Score 368.4; DB 9; Length 1168;
Best Local Similarity 99.7%; Pred. No. 2.8e-87;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 123 GTTGGCTCTGGGTCGGTTGAGGGCGGCAGTCTCGSCATGTAGTTCTCAGGTTCTGG 182
Db |||||
QY 871 GCTGGCTCTGGGTCGGTTGAGGGCGGCAGTCTCGSCATGTAGTTCTCAGGTTCTGG 812
Db |||||
QY 183 CAGGGCTGCAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGCT 242
Db |||||
QY 811 CAGGGCTGCAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGCT 752
Db |||||
QY 243 GCACCAAGCTACCAAGTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGGCTCACC' 302
Db |||||
QY 751 GCACCAAGCTACCAAGTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGGCTCACC 692
Db |||||
QY 303 TGACACAGGGCTAGACCCCTCAGAGTGGCGCCGCTGTCCTGAGTGTGGGTGCTGCT 362
Db |||||
QY 691 TGACACAGGGCTAGACCCCTCAGAGTGGCGCCGCTGTCCTGAGTGTGGGTGCTGCT 632
Db |||||
QY 363 CAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCAGGAGCACCCCTGG 422
Db |||||
QY 631 CAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCAGGAGCACCCCTGG 572
Db |||||
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCGGAAGAGGGCTCGAGACAC 482

Db 571 CCTGGGTTCCATGCCCGCTCCAGCTGTGGCCCTCGTCCCGAAGAAGGGCTCGAGACAC 512
QY 483 CCGATCGTGA 492
Db 511 CCGATCGTGA 502
RESULT 14
ADJ56254/c
ID ADJ56254 standard; cDNA; 1168 BP.
XX
AC ADJ56254;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human cDNA differentially expressed in MYCN activated cells SeqID 60.
XX
KW human; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
OS Homo sapiens.
XX
PN US2003119009-A1.
XX
PD 26-JUN-2003.
XX
PF 25-FEB-2002; 2002US-00084817.
XX
PR 23-FEB-2001; 2001US-0270784P.
XX
PA (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHE J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
DR WPI; 2003-635698/60.
XX
PT New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 60; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 1168 BP; 249 A; 346 C; 339 G; 234 T; 0 U; 0 Other;
Query Match 74.9%; Score 368.4; DB 10; Length 1168;
Best Local Similarity 99.7%; Pred. No. 2.8e-87;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 123 GTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCCGSCATGTAGTGTCTCAGGTTCTGG 182

Db 871 GCTGGCTCTGGTCTGGTTGGAGGGCGCAGTGTCTGGCATGTAGTGTCTCAGGTTCTGG 812
QY 183 CAGGGCTGCAGAAAGGCTGTTCAGCTGCGCTGAGATTGTCAAGCGCGGCTCCAGGCCT 242
Db 811 CAGGGCTGCAGAAAGGCTGTTCAGCTGCGCTGAGATTGTCAAGCGCGGCTCCAGGCCT 752
QY 243 GCACCAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACC 302
Db 751 GCACCAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCACC 692
QY 303 TGACACAGGGCTAGACCCCTCACAGTGGCGCCGCCATGTGCCTGACGTGTGGGTGCTGCT 362
Db 691 TGACACAGGGCTAGACCCCTCACAGTGGCGCCGCCATGTGCCTGACGTGTGGGTGCTGCT 632
QY 363 CAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCTGG 422
Db 631 CAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCTGG 572
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGACAC 482
Db 571 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGACAC 512
QY 483 CCGATCGTGA 492
Db 511 CCGATCGTGA 502
RESULT 15
AAC03605
ID AAC03605 standard; cDNA; 303 BP.
XX
AC AAC03605;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3603.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG03599.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 3603; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX
SQ Sequence 303 BP; 57 A; 81 C; 105 G; 60 T; 0 U; 0 Other;

Query Match 48.0%; Score 236; DB 3; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.8e-52;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
68 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 127
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
128 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 187
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
188 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 247
QY 181 GGCAGGGCTGCAGGAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGCGGTTCCC 236
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
248 GGCAGGGCTGCAGGAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGCGGTTCCC 303

Search completed: January 31, 2005, 16:45:02
Job time : 395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:06:38 ; Search time 135.925 Seconds
(without alignments)
433.254 Million cell updates/sec

Title: US-10-057-813-14
Perfect score: 861
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....SMPSSSCGPRRRRRRTS 163

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	163	9	US-09-731-872-459 Sequence 459, App
2	861	100.0	163	10	US-09-876-997-459 Sequence 459, App
3	813	94.4	163	11	US-09-978-360A-668 Sequence 668, App
4	305	35.4	72	17	US-10-425-115-237341 Sequence 237341,
5	227	26.4	281	15	US-10-424-599-237821 Sequence 237821,
6	224	26.0	242	17	US-10-425-115-254668 Sequence 254668,
7	204.5	23.8	245	15	US-10-424-599-229595 Sequence 229595,
8	203.5	23.6	276	17	US-10-425-115-308895 Sequence 308895,
9	203.5	23.6	276	17	US-10-425-115-308929 Sequence 308929,
10	203.5	23.6	278	15	US-10-425-114-60670 Sequence 60670, A
11	203.5	23.6	301	15	US-10-425-114-61107 Sequence 61107, A
12	203.5	23.6	315	15	US-10-425-114-54515 Sequence 54515, A
13	203.5	23.6	315	15	US-10-425-114-58679 Sequence 58679, A

14	203.5	23.6	325	15	US-10-425-114-60913 Sequence 60913, A
15	202	23.5	178	17	US-10-425-115-232929 Sequence 232929,
16	202	23.5	671	16	US-10-437-963-109378 Sequence 109378,
17	201.5	23.4	269	16	US-10-437-963-156973 Sequence 156973,
18	200.5	23.3	263	16	US-10-437-963-152741 Sequence 152741,
19	200	23.2	262	16	US-10-437-963-191078 Sequence 191078,
20	199.5	23.2	275	16	US-10-767-701-45240 Sequence 45240, A
21	198	23.0	294	15	US-10-425-114-60641 Sequence 60641, A
22	193.5	22.5	261	17	US-10-425-115-308917 Sequence 308917,
23	191	22.2	265	17	US-10-425-115-201476 Sequence 201476,
24	190.5	22.1	261	16	US-10-767-701-42504 Sequence 42504, A
25	188	21.8	287	16	US-10-437-963-152743 Sequence 152743,
26	188	21.8	359	17	US-10-425-115-231292 Sequence 231292,
27	187	21.7	261	16	US-10-437-963-184232 Sequence 184232,
28	185	21.5	320	17	US-10-739-930-10929 Sequence 10929, A
29	172	20.0	318	17	US-10-425-115-266917 Sequence 266917,
30	102.5	11.9	242	16	US-10-437-963-134549 Sequence 134549,
31	96	11.1	198	17	US-10-425-115-365780 Sequence 365780,
32	88.5	10.3	213	17	US-10-425-115-278508 Sequence 278508,
33	88	10.2	255	17	US-10-425-115-342952 Sequence 342952,
34	88	10.2	581	16	US-10-437-963-201766 Sequence 201766,
35	87.5	10.2	625	14	US-10-156-761-10693 Sequence 10693, A
36	86.5	10.0	174	14	US-10-094-749-2987 Sequence 2987, Ap
37	86.5	10.0	353	17	US-10-425-115-224520 Sequence 224520,
38	86.5	10.0	359	15	US-10-296-115-1274 Sequence 1274, Ap
39	86.5	10.0	999	16	US-10-437-963-181440 Sequence 181440,
40	86	10.0	186	17	US-10-425-115-220639 Sequence 220639,
41	85.5	9.9	639	16	US-10-408-765A-1054 Sequence 1054, Ap
42	85.5	9.9	774	17	US-10-723-860-1706 Sequence 1706, Ap
43	85	9.9	436	14	US-10-156-761-11365 Sequence 11365, A
44	84.5	9.8	234	15	US-10-108-260A-2464 Sequence 2464, Ap
45	84.5	9.8	283	16	US-10-437-963-137561 Sequence 137561,

ALIGNMENTS

RESULT 1

US-09-731-872-459
; Sequence 459, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-731-872-459

Query Match 100.0%; Score 861; DB 9; Length 163;
Best Local Similarity 100.0%; Pred. No. 9.7e-77;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS	60
Db	1	MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS	60
QY	61	GRAAGKAVSCAEIVKRRVPGHLQTLKLRFLQTEDSWVPASPDITGLDPLTVRRHVPVAVWL	120
Db	61	GRAAGKAVSCAEIVKRRVPGHLQTLKLRFLQTEDSWVPASPDITGLDPLTVRRHVPVAVWL	120

QY 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRRRRRARDTRS 163
Db 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRRRRRARDTRS 163

RESULT 2
US-09-876-997-459
; Sequence 459, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-997-459

Query Match 100.0%; Score 861; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 9.7e-77;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120

QY 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRRRRRARDTRS 163
Db 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRRRRRARDTRS 163

RESULT 3
US-09-978-360A-668
; Sequence 668, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273

; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 668
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52..-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)
; OTHER INFORMATION: unknown
US-09-978-360A-668

Query Match 94.4%; Score 813; DB 11; Length 163;
Best Local Similarity 94.5%; Pred. No. 5.3e-72;
Matches 154; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
QY 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRRRRRARDTRS 163
Db 121 LXRDLDPNECGYQPPGAPPGGLGSMPPSSSCGPRSRKXAXXTRS 163

RESULT 4
US-10-425-115-237341
; Sequence 237341, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 237341
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_148039C.1.pep
US-10-425-115-237341

Query Match 35.4%; Score 305; DB 17; Length 72;
Best Local Similarity 86.1%; Pred. No. 2.3e-22;
Matches 62; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLAQRLEGGSALQVVLGS 60

QY 61 GRAAGKAVSCAE 72
||| ||| ||| |||
Db 61 GRAARNVSVFAE 72

RESULT 5

US-10-424-599-237821
; Sequence 237821, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 237821
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56778C.1.pap
US-10-424-599-237821

Query Match	26.4%;	Score 227;	DB 15;	Length 281;
Best Local Similarity	31.8%;	Pred. No. 6.2e-14;		
Matches 56;	Conservative 24;	Mismatches 66;	Indels 30;	Gaps 4;

QY	1	MEHYRKAGSV	ELP	APSP	MPQ	LP	PD	TL	EM	R	VR	D	G	S	K	I	R	N	L	L	G	L	A	L	R	L	E	G	S	A	R	H	V	V	F	S	G	S	60																					
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																					
Db	33	MDRYQK	-----	VEK	P	R	A	E	T	P	I	D	E	N	I	R	I	T	S	Q	G	R	M	R	Y	I	T	Y	A	M	S	L	L	Q	E	K	S	N	E	I	V	F	K	A	M	85														
QY	61	G	R	A	A	G	K	V	S	C	A	E	I	V	K	R	R	V	P	G	L	H	Q	L	T	K	L	R	F	L	Q	T	E	D	S	W	P	A	S	P	D	T	G	L	D	P	L	T	V	R	R	H	V	P	A	V	W	V	L	120
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:							
Db	86	G	R	A	I	N	K	T	V	T	I	V	E	L	I	K	R	I	V	G	L	H	O	N	T	A	I	S	T	D	I	T	D	T	W	E	P	L	--	E	E	G	L	L	P	L	E	T	R	H	V	S	M	I	T	V	143			

Qy	121	LSRDPLDPNECGYQP-----GAPPGISMPSSSCGPRRRAR	159
	:	:	:
Db	144	LSKKELTSSVGYQPPLPADQVKAAATDFDYEGESPNG----RAHGRGGRGGRGCR	195
	:	:	:

RESULT 6

US-10-425-115-254668
; Sequence 254668, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 254668
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(242)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_163838C.1.pep
US-10-425-115-254668

Query Match 26.0%; Score 224; DB 17; Length 242;

Best Local Similarity	33.1%;	Pred. No. 1e-13;							
Matches	57;	Conservative	23;	Mismatches	70;	Indels	22;	Gaps	5;

[illegible]

RESULT 7

US-10-424-599-229595
; Sequence 229595, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229595
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49349C.1.pep
US-10-424-599-229595

Query Match	23.8%	Score 204.5	DB 15	Length 245
-------------	-------	-------------	-------	------------

[illegible]

RESULT 8

```

US-10-425-115-308895
; Sequence 308895, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 308895
; LENGTH: 276

```



```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144024C.1.pap
US-10-425-115-232929

Query Match      23.5%  Score 202; DB 17; Length 178;
Best Local Similarity 30.1%; Pred. No. 1e-11;
Matches 53; Conservative 23; Mismatches 70; Indels 30; Gaps 4;

QY 1 MEHYRKAGSVLPAPSPMPQLPDTLEMRVRDGSKIRNLGLALGRLEGGSARHVVVPSGS 60
Db 1 MDRYQR-----VEKPREESPIGANEIRITAQGRPRNYITYALALLQDNATDDIVIKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVL 120
Db 54 GRAINKTVAIVELLKRRRIAGLHQNTSIESIDITDTWEPL--EEGLVTLETIRHVSLLITIK 111

QY 121 LSRDPLDPNECGYQPP-----GAPGLGSMPSSSCGPRRRRAR 159
Db 112 LSKKELDTSSPGYQPPIPADQVRPAADFQDAEAVPSGRG----RGRGRRSRGRGR 163
```

Search completed: January 26, 2005, 15:28:35
Job time: 136.925 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:07:24 ; Search time 2520 Seconds
(without alignments)
9232.751 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgcatcgtga 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

ALIGNMENTS

ALIGNMENTS

RESULT 1

BC032136

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	492	100.0	871	9	BC032136	BC032136 Homo sapi
2	492	100.0	199114	9	AL160270	AL160270 Human DNA
3	484	98.4	884	6	BD139494	BD139494 Extended
4	374.4	76.1	461	6	AR412580	AR412580 Sequence
5	374.4	76.1	461	6	AX969414	AX969414 Sequence
6	374.4	76.1	461	6	BD108133	BD108133 EST and e
7	372	75.6	238811	2	AC098197	AC098197 Rattus no
8	372	75.6	242324	2	AC110351	AC110351 Rattus no
9	370.4	75.3	834	10	MMU293897	AJ293897 Mus muscu
10	370.4	75.3	863	10	BC047068	BC047068 Mus muscu
11	370.4	75.3	168268	10	AL807796	AL807796 Mouse DNA
12	236	48.0	303	6	AX887740	AX887740 Sequence
13	236	48.0	303	6	BD027350	BD027350 Sequence
14	236	48.0	305	6	BD077019	BD077019 5' EST of
15	85.4	17.4	1554	5	BC076268	BC076268 Danio rer
16	85.4	17.4	189318	2	CR376821	CR376821 Danio rer
17	85.4	17.4	213678	5	BX005012	BX005012 Zebrafish
18	81.6	16.6	242077	2	AC095805	AC095805 Rattus no
19	81.6	16.6	251145	2	AC108546	AC108546 Rattus no

20	77.8	15.8	1361	10	BC016085	BC016085 Mus muscu
21	74.6	15.2	981	6	CQ735345	CQ735345 Sequence
22	74.6	15.2	1183	9	AY034074	AY034074 Homo sapi
23	74.6	15.2	1473	9	BC007270	BC007270 Homo sapi
24	74.6	15.2	1483	9	BC002497	BC002497 Homo sapi
25	74.6	15.2	1941	6	AR338514	AR338514 Sequence
26	74.6	15.2	2375	6	AR447527	AR447527 Sequence
c	74.6	15.2	174559	9	AC015720	AC015720 Homo sapi
28	73	14.8	2015	9	AK000381	AK000381 Homo sapi
29	59.2	12.0	913	5	BX931336	BX931336 Gallus ga
30	55	11.2	125020	9	AF429315	AF429315 Homo sapi
c	55	11.2	167543	2	AC107974	AC107974 Homo sapi
32	53.2	10.8	1242	5	BX934615	BX934615 Gallus ga
c	48.2	9.8	65476	2	AC023362	AC023362 Mus muscu
34	46	9.3	799	3	BT011363	BT011363 Drosophil
c	45.4	9.2	110000	2	LMFLCHR34_13	Continuation (14 o
36	44.4	9.0	648	6	CQ592263	CQ592263 Sequence
37	44.4	9.0	2000	6	AX655393	AX655393 Sequence
c	43.4	8.8	125020	9	AF429315	AF429315 Homo sapi
39	42.8	8.7	882	3	AY089619	AY089619 Drosophil
c	41	8.3	2000	6	AX655393	AX655393 Sequence
c	41	8.3	182559	9	AC073476	AC073476 Homo sapi
42	40.6	8.3	231739	9	AC011811	AC011811 Homo sapi
c	40.4	8.2	244813	2	AC097868	AC097868 Rattus no
44	40.4	8.2	275726	2	AC118128	AC118128 Rattus no
c	40.4	8.2	300327	1	AE017228	AE017228 Mycobacte

BC032136 871 bp mRNA linear PRI 30-JUN-2004
Homo sapiens chromosome 9 open reading frame 23, transcript variant 1, mRNA (cdna clone MGC:29635 IMAGE:4869926), complete cds.

BC032136

BC032136.2 GI:33879663

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 871)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 871)

Strausberg,R.

Direct Submission

Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:21618954.
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Iaric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: h Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22325369.

FEATURES
source

1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:29635 IMAGE:4869926"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH MGC_49"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

gene

1..871
/gene="C9orf23"
/note="synonyms: MGC29635, ba296L22.5"
/db_xref="LocusID:138716"
53..544
/gene="C9orf23"
/codon_start=1
/product="chromosome 9 open reading frame 23 protein"
/protein_id="AAH32136.1"
/db_xref="GI:21618955"
/db_xref="LocusID:138716"
/translation="MEHYRKAGSVELPAPSPMPQLPPDLEMRVDRGSKIRNLLGLAL
GRLEGGSARHVVFGSGRAAGKAVSCAIEIVKRRVPLGLHQLTKLFLQTEDSWVPASPD
TGLDPLTVRRHVPAVWVLLSRDPLDNECYQPPGAPPLGLSMPSSCGPRRRARD
TRS"

CDS

ORIGIN

Query Match 100.0%; Score 492; DB 9; Length 871;
Best Local Similarity 100.0%; Pred. No. 1.4e-114;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
53 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 112
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 CTACCTCCTGATACCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 CTACCTCCTGATACCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 172
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCT 232
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 GGCAGGGCTGAGGAAAGGCTGTACGCTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
233 GGCAGGGCTGAGGAAAGGCTGTACGCTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC 292

QY 241 CTGCACCAGCTCACCAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
293 CTGCACCAGCTCACCAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 352
QY 301 CCTGACACAGGGCTAGACCCCTCCTACAGTGCCTGCGCCATGTGCTGCTGAGTGTGGTGTG 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
353 CCTGACACAGGGCTAGACCCCTCCTCAGTGCCTGCGCCATGTGCTGCTGAGTGTGGTGTG 412
QY 361 CTCAGCCGGGACCCCTTGAGCCCAATGAGTGTGTTACCAACCCAGGAGCACCCCT 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
413 CTCAGCCGGGACCCCTTGAGCCCAATGAGTGTGTTACCAACCCAGGAGCACCCCT 472
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
473 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 532
QY 481 ACCCGATCGTGA 492
Db |||||||||||||||
533 ACCCGATCGTGA 544

RESULT 2
AL160270/c
LOCUS

DEFINITION

AL160270 199114 bp DNA linear PRI 27-APR-2001
Human DNA sequence from clone RP11-296L22 on chromosome 9, complete
sequence.

ACCESSION

AL160270.19 GI:13443330

VERSION

HTG.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 199114)

REFERENCE

Sehra,H.

AUTHORS

Direct Submission

TITLE

Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgehire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

JOURNAL

On Mar 24, 2001 this sequence version replaced gi:13274314.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

COMMENT

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Swi,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
RP11-296L22 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6

FEATURES

source

1..199114
/organism="Homo sapiens"
/mol_type="genomic DNA"

repeat_region /db_xref="taxon:9606"
repeat_region /chromosome="9"
repeat_region /clone="RP11-296L22"
repeat_region /clone_lib="RPCI-11.2"
repeat_region 632. .854
repeat_region /note="AluJo repeat: matches 85. .312 of consensus"
repeat_region 855. .1006
repeat_region /note="38 copies 4 mer ggaa 80% conserved"
repeat_region 860. .1003
repeat_region /note="2 copies 72 mer 84% conserved"
repeat_region 862. .1003
repeat_region /note="2 copies 71 mer 83% conserved"
repeat_region 1135. .1402
repeat_region /note="AluSx repeat: matches 1. .294 of consensus"
repeat_region 1468. .1644
repeat_region /note="FAM repeat: matches 1. .175 of consensus"
repeat_region 1654. .1709
repeat_region /note="4 copies 14 mer 83% conserved"
repeat_region 1719. .1999
repeat_region /note="AluSx repeat: matches 1. .303 of consensus"
repeat_region 2039. .2173
repeat_region /note="L1P2 repeat: matches 5685. .5817 of consensus"
repeat_region 3441. .3757
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 3767. .3826
repeat_region /note="L1MB4 repeat: matches 6131. .6183 of consensus"
repeat_region 5346. .5626
repeat_region /note="AluJb repeat: matches 1. .276 of consensus"
repeat_region 5627. .5696
repeat_region /note="5 copies 14 mer 78% conserved"
repeat_region 5663. .5698
repeat_region /note="18 copies 2 mer ga 100% conserved"
repeat_region 6509. .6604
repeat_region /note="MER5A repeat: matches 10. .109 of consensus"
repeat_region 6869. .7039
repeat_region /note="MIR repeat: matches 73. .257 of consensus"
repeat_region 7474. .7536
repeat_region /note="L2 repeat: matches 2634. .2693 of consensus"
repeat_region 7537. .7829
repeat_region /note="AluJb repeat: matches 1. .305 of consensus"
repeat_region 7830. .8610
repeat_region /note="L2 repeat: matches 1814. .2634 of consensus"
repeat_region 8689. .8789
repeat_region /note="L2 repeat: matches 1432. .1538 of consensus"
repeat_region 9468. .9757
repeat_region /note="AluJb repeat: matches 1. .294 of consensus"
repeat_region 9859. .10157
repeat_region /note="AluSx repeat: matches 8. .307 of consensus"
repeat_region 11656. .11741
repeat_region /note="L2 repeat: matches 2250. .2339 of consensus"
repeat_region 11831. .12045
repeat_region /note="AluJb repeat: matches 85. .294 of consensus"
repeat_region 14904. .15040
repeat_region /note="FLAM C repeat: matches 5. .141 of consensus"
repeat_region 15048. .15319
repeat_region /note="AluY repeat: matches 40. .311 of consensus"
repeat_region 15327. .15522
repeat_region /note="AluJb repeat: matches 118. .310 of consensus"
repeat_region 16010. .16045
repeat_region /note="9 copies 4 mer caca 86% conserved"
repeat_region 16247. .16381
repeat_region /note="FLAM A repeat: matches 1. .132 of consensus"
repeat_region 16575. .16715
repeat_region /note="MIR repeat: matches 2. .160 of consensus"
repeat_region 16955. .17374
repeat_region /note="L2 repeat: matches 1273. .1754 of consensus"
repeat_region 17525. .17650
repeat_region /note="AluJb repeat: matches 1. .126 of consensus"
repeat_region 17722. .17922
repeat_region /note="L1MD3 repeat: matches 6867. .7071 of consensus"
repeat_region 18044. .18366
repeat_region /note="L2 repeat: matches 2013. .2338 of consensus"
repeat_region 18574. .18883

repeat_region /note="AluY repeat: matches 1. .310 of consensus"
repeat_region 19055. .19190
repeat_region /note="MIR repeat: matches 70. .216 of consensus"
repeat_region 19465. .19773
repeat_region /note="AluSx repeat: matches 1. .308 of consensus"
repeat_region 20467. .20757
repeat_region /note="AluSx repeat: matches 1. .292 of consensus"
repeat_region 20899. .21198
repeat_region /note="AluJo repeat: matches 1. .297 of consensus"
repeat_region 21374. .21565
repeat_region /note="MIR repeat: matches 14. .228 of consensus"
repeat_region 22198. .22393
repeat_region /note="MIR repeat: matches 2. .241 of consensus"
repeat_region 22396. .23932
repeat_region /note="PTR5 repeat: matches 145. .2438 of consensus"
repeat_region 24018. .24527
repeat_region /note="L2 repeat: matches 2181. .2710 of consensus"
repeat_region 25222. .25521
repeat_region /note="AluY repeat: matches 1. .299 of consensus"
repeat_region 25522. .25685
repeat_region /note="FAM repeat: matches 1. .168 of consensus"
repeat_region 26380. .26655
repeat_region /note="AluJb repeat: matches 21. .303 of consensus"
repeat_region 26817. .27055
repeat_region /note="L1MC5 repeat: matches 7302. .7574 of consensus"
repeat_region 27056. .27363
repeat_region /note="AluSx repeat: matches 1. .308 of consensus"
repeat_region 27364. .27544
repeat_region /note="L1MC5 repeat: matches 7574. .7744 of consensus"
repeat_region 27601. .27730
repeat_region /note="L1MC5 repeat: matches 7725. .7861 of consensus"
repeat_region 29389. .29688
repeat_region /note="L2 repeat: matches 2430. .2710 of consensus"
repeat_region 30074. .30156
repeat_region /note="L2 repeat: matches 2417. .2499 of consensus"
repeat_region 30546. .30841
repeat_region /note="AluSx repeat: matches 1. .298 of consensus"
repeat_region 32521. .32823
repeat_region /note="AluJo repeat: matches 1. .304 of consensus"
repeat_region 33433. .33653
repeat_region /note="L1PB2 repeat: matches 5927. .6155 of consensus"
repeat_region 33705. .33932
repeat_region /note="AluJb repeat: matches 1. .225 of consensus"
repeat_region 34683. .34976
repeat_region /note="AluJb repeat: matches 1. .297 of consensus"
repeat_region 35077. .35368
repeat_region /note="AluSg1 repeat: matches 1. .293 of consensus"
repeat_region 37049. .37547
repeat_region /note="L2 repeat: matches 1855. .2412 of consensus"
repeat_region 37618. .37671
repeat_region /note="L2 repeat: matches 2691. .2744 of consensus"
repeat_region 38809. .39109
repeat_region /note="AluSc repeat: matches 1. .300 of consensus"
repeat_region 39720. .39836
repeat_region /note="L2 repeat: matches 2592. .2706 of consensus"
repeat_region 40119. .40338
repeat_region /note="AluJb repeat: matches 82. .301 of consensus"
repeat_region 40564. .40863
repeat_region /note="AluSx repeat: matches 1. .300 of consensus"
repeat_region 42031. .43547
repeat_region /note="CpG island"
repeat_region /evidence=not_experimental
repeat_region 42904. .42970
repeat_region /note="Sequence from overlapping clone BA556O4 (AL359088).
Assembly confirmed by restriction digest."
repeat_region 42917. .42976
repeat_region /note="10 copies 6 mer ggcgcg 71% conserved"
repeat_region 44087. .44398
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 45303. .45481
repeat_region /note="AluSq repeat: matches 135. .313 of consensus"
repeat_region 45482. .45794
repeat_region /note="AluY repeat: matches 1. .311 of consensus"


```
repeat_region 45795..45925
/note="AluSg repeat: matches 5. .135 of consensus"
repeat_region 45953..46182
/note="L1ME2 repeat: matches 5933. .6155 of consensus"

Query Match 100.0%; Score 492; DB 9; Length 199114;
Best Local Similarity 100.0%; Pred. No. 8.6e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db 195948 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 195889

QY 61 CTACCTCCTGATACCTTGAGATGCGGTCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 195888 CTACCTCCTGATACCTTGAGATGCGGTCGAGATGGCAGCAAAATTGCAACCTGCTG 195829

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 195828 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 195769

QY 181 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db 195768 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGCGGTCCCAGGC 195709

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 195708 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 195649

QY 301 CCTGACACAGGGCTAGACCCCTCAACAGTGGCGGCCCATGTGCCCTGCAGTGTGGTGCTG 360
Db 195648 CCTGACACAGGGCTAGACCCCTCAACAGTGGCGGCCCATGTGCCCTGCAGTGTGGTGCTG 195589

QY 361 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 195588 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 195529

QY 421 GGCTGGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 195528 GGCTGGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 195469

QY 481 ACCCGATCGTGA 492
Db 195468 ACCCGATCGTGA 195457

RESULT 3
BD139494
LOCUS BD139494
DEFINITION Extended cDNA of secretory protein.
ACCESSION BD139494
VERSION BD139494.1 GI:23234439
KEYWORDS JP 2002508182-A/246.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
Bougueret,L., Duclert,A. and Edwards,J.B.D.M.
Extended cDNA of secretory protein
Patent: JP 2002508182-A 246 19-MAR-2002;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2002508182-A/246
PD 19-MAR-2002
PF 17-DEC-1998 JP 2000539136
PR 17-DEC-1997 US 60/069957,09-FEB-1998 US 60/074121 PR
13-APR-1998 US 60/081563,10-AUG-1998 US 60/096116 PI LYDIE
BOUGUERET,AYMERIC DUCLERT,JEAN BAPTISTE DUMAS MILNE PI EDWARDS
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC
Von Heijne matrix
```

```
CC score 4.69999980926514
CC seq LGLALGRLEGGSA/RH
FH Key Location/Qualifiers
FT CDS 69..557
FT sig_peptide 69..224
FT polyA_signal 849..854
FT polyA_site 870..883.
FEATURES
Location/Qualifiers
source
1..884
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.4%; Score 484; DB 6; Length 884;
Best Local Similarity 97.8%; Pred. No. 1.6e-112;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db 69 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 128

QY 61 CTACCTCCTGATACCTTGAGATGCGGTCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 129 CTACCTCCTGATACCTTGAGATGCGGTCGAGATGGCAGCAAAATTGCAACCTGCTG 188

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 189 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 248

QY 181 GGAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db 249 GGAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGCGGTCCCAGGC 308

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 309 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 368

QY 301 CCTGACACAGGGCTAGACCCCTCAACAGTGGCGGCCCATGTGCCCTGCAGTGTGGTGCTG 360
Db 369 CCTGACACAGGGCTAGACCCCTCAACAGTGGCGGCCCATGTGCCCTGCAGTGTGGTGCTG 428

QY 361 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 429 CTCAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCGAGGACACCCCT 488

QY 421 GGCTGGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 489 GGCTGGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGARAC 548

QY 481 ACCCGATCGTGA 492
Db 549 ACCCGATCGTGA 560
```

```
RESULT 4
AR412580
LOCUS AR412580
DEFINITION Sequence 217 from patent US 6639063.
ACCESSION AR412580
VERSION AR412580.1 GI:40167690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 461)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 217 28-OCT-2003;
FEATURES
Location/Qualifiers
source
1..461
/organism="unknown"
/mol_type="genomic DNA"
```

ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;
Best Local Similarity 98.4%; Pred. No. 1.3e-84;
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
|||
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 137
|||
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
|||
Db 138 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
|||
QY 121 GGGTTGGCTCTGGTTCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 180
|||
Db 198 GGGTTGGCTCTGGTTCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 257
|||
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 240
|||
Db 258 GGCAGGGCTGCAGGAAAGCTGTGAGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 317
|||
QY 241 CTGCACCAAGCTACCGTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 300
|||
Db 318 CTGCACCAAGCTACCGTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 377
|||
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCTGAGGACAGCTGGGTCCAGCCTCA 360
|||
Db 378 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCTGAGGACAGCTGGGTCCAGCCTCA 437
|||
QY 361 CTCAGCCGGGACCCCTGGACC 382
|||
Db 438 CTCAGCCGGGDMCCCTGGACCC 459
|||

RESULT 5

AX969414
LOCUS AX969414 461 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 217 from Patent EP1104808.
ACCESSION AX969414
VERSION AX969414.1 GI:40975747
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 217 06-JUN-2001;
Genset (FR)

FEATURES

source Location/Qualifiers
1. .461
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
78. .>461
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF14522.1"
/db_xref="GI:40975748"
/translation="MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVDRGSKIRNLLGLAL
GRLEGGSARHVFPVFGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPD
TGLDPLTVRRHVPVAVVLLSRXPWTQ"
sig_peptide 78. .233
/note="Von Heijne matrix score 4.69999980926514 seq
LGLALGRLEGGSA/RH"

ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;
Best Local Similarity 98.4%; Pred. No. 1.3e-84;
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
|||
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 137
|||
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
|||
Db 138 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
|||
QY 121 GGGTTGGCTCTGGTTCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 180
|||
Db 198 GGGTTGGCTCTGGTTCGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 257
|||
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 240
|||
Db 258 GGCAGGGCTGCAGGAAAGCTGTGAGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 317
|||
QY 241 CTGCACCAAGCTACCGTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 300
|||
Db 318 CTGCACCAAGCTACCGTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 377
|||
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCTGAGGACAGCTGGGTCCAGCCTCA 360
|||
Db 378 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCTGAGGACAGCTGGGTCCAGCCTCA 437
|||
QY 361 CTCAGCCGGGACCCCTGGACC 382
|||
Db 438 CTCAGCCGGGDMCCCTGGACCC 459
|||

RESULT 6

BD108133
LOCUS BD108133 461 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD108133
VERSION BD108133.1 GI:23202951
KEYWORDS JP 2002010789-A/210.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 210 15-JAN-2002;
GENSET CORP

COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/210
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 4.69999980926514
CC seq LGLALGRLEGGSA/RH
FH Key Location/Qualifiers
FT CDS 78. .461
FT sig_peptide 78. .233.
FEATURES Location/Qualifiers
1. .461
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 76.1%; Score 374.4; DB 6; Length 461;
Best Local Similarity 98.4%; Pred. No. 1.3e-84;
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
|||||
Db 78 ATGGAGCACTACCGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 137
|||||
QY 61 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
|||||
Db 138 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
|||||
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGCGCGCAGTCTCGGCATGTAGTCTCAGGTTCT 180
|||||
Db 198 GGGTTGGCTCTGGTTCGGTTGGAGGCGCGCAGTCTCGGCATGTAGTCTCAGGTTCT 257
|||||
QY 181 GGCAGGGCTGCAGGAAGCTGTCAGCTGCGCTGAGATTGTCAAGCGCGGTCCCAGGC 240
|||||
Db 258 GGCAGGGCTGCAGGAAGCTGTCAGCTGCGCTGAGATTGTCAAGCGCGGTCCCAGGC 317
|||||
QY 241 CTGCACCAGCTACCAAGCTACGTTTCCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
|||||
Db 318 CTGCACCAGCTACCAAGCTACGTTTCCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 377
|||||
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGCAGTGTGGTGCTG 360
|||||
Db 378 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGCAGTGTGGTGCTG 437
|||||
QY 361 CTCAGCGGGACCCCTGGACC 382
|||||:|
Db 438 CTCAGCGGGDMCCCTGGACCC 459

RESULT 7

AC098197/c
LOCUS
DEFINITION
AC098197
AC098197.8 GI:30581495
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 238811)

REFERENCE

AUTHORS

Muzny,D., Marie., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allien,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., Mcleod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemeleh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 238811)

Worley,K.C.

Direct Submission

Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 238811)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 13, 2003 this sequence version replaced gi:24819301.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GIKF

Center clone name: CH230-9714

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 228248 bases at least Q40

Consensus quality: 230539 bases at least Q30

Consensus quality: 232099 bases at least Q20

Estimated insert size: 236470; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 44755: contig of 44755 bp in length

* 44756 44855: gap of unknown length
* 44856 129903: contig of 85048 bp in length
* 129904 130003: gap of unknown length
* 130004 227138: contig of 97135 bp in length
* 227139 227238: gap of unknown length
* 227239 228983: contig of 1745 bp in length
* 228984 229083: gap of unknown length
* 229084 230325: contig of 1242 bp in length
* 230326 230425: gap of unknown length
* 230426 232444: contig of 2019 bp in length
* 232445 232544: gap of unknown length
* 232545 234255: contig of 1711 bp in length
* 234256 234355: gap of unknown length
* 234356 235710: contig of 1355 bp in length
* 235711 235810: gap of unknown length
* 235811 236820: contig of 1010 bp in length
* 236821 236920: gap of unknown length
* 236921 238811: contig of 1891 bp in length.

FEATURES

source
1. .238811
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9714"
3538. .4289
/note="clone boundary
clone_end:Sp6
site:ECORI
end sequence:BH282839"
44856. .46019
/note="wgs contig"
128441. .129903
/note="wgs_contig"

misc_feature
misc_feature
misc_feature

ORIGIN

Query Match 75.6%; Score 372; DB 2; Length 238811;
Best Local Similarity 84.8%; Pred. No. 2.9e-84;
Matches 417; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 60
Db 150054 ATGGAGCACTACCGGAGGCTGGCTCTGTAGAGCTCCAGCCTCATCACCATGCCCCAG 149995

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 149994 CTACCTCCCGACACGCTGGAATCGGGTCCGAGATGGGAGCAAAATCCGAAACCTACTA 149935

QY 121 GGGTTGGCTCTGGGTCGGTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 149934 GGGCTGGCACTGGCTCGTTTGAAGGAGGAAGCACAAAGGCACGTGGTGTCTCAGGCTCC 149875

QY 181 GGCAGGGCTGCAGGAAGGCTGTGCTGAGTGCCTGAGATTGTCAAGCGGCGGFTCCAGGC 240
Db 149874 GGC CGGGCCGCTGGAAGGCGCTCAGCTGTGCAGAGATTGTCAAACGGCGGTTCCGGGC 149815

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGSGTCCAGCCTCA 300
Db 149814 CTACACCAAGCTACCAAGCTTCGCTTCTCTGCAAACTGAGGACAGCTGGGTCCCAACTTCA 149755

QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGC CGCCGCATGTGCCTGCAGTGTGGTGCTG 360
Db 149754 CCGGACACGGGCTTGACCCCTTCACAGTCCGACGCCATGTGCCTGCAGTGTGGTACTG 149695

QY 361 CTCAGCCGGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCTCT 420
Db 149694 CTTAGTCGGGACCCCTTGACCCCAAGTGAATGTGGCTACCAACCCCGAGTGCACCTCCT 149635

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTC CGGAAGAGGCTCGAGAC 480
Db 149634 GGCCTGGGCTCCATACCTAGTTCAGCTGTGGTCCAGACCCCGAAGGAGGCTCGGGAC 149575

QY 481 ACCCGATCGTGA 492

Db 149574 ACCCGGTCCTGA 149563

RESULT 8

AC110351/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC110351 242324 bp DNA linear HTG 12-OCT-2002
Rattus norvegicus clone CH230-217H18, WORKING DRAFT SEQUENCE.
AC110351
AC110351.4 GI:23915435
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 242324)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemeleh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 242324)

Worley,K.C.

Direct Submission

Submitted (11-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 242324)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Oct 12, 2002 this sequence version replaced gi:21739791. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRLK
Center clone name: CH230-217H18
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 22326 bases at least Q40
Consensus quality: 22424 bases at least Q30
Consensus quality: 225710 bases at least Q20
Estimated insert size: 228121; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 242324: contig of 242324 bp in length.

FEATURES
Source
Location/Qualifiers
1. .242324
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-217H18"
1. .1109
/note="wgs_end_extension
clone_end:T7"
1611. .3024
/note="wgs_end_extension
clone_end:T7"
4338. .42885
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:RWBMD45TV"
45628. .45664
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:RWBMD45TV"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match 75.6%; Score 372; DB 2; Length 242324;
Best Local Similarity 84.8%; Pred. No. 2.9e-84;

Matches 417; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 60
|||||
Db 13008 ATGGAGCACTACCGGAGGCTGGCTCTGTAGAGCTCCAGCCTCATCACCATGCCCCAG 12949
|||||

QY 61 CTACCTCTGATACCCCTTGAGATGGGGTCCGAGATGGCAGCAAAATTCGAACCTGCTG 120
|||||
Db 12948 CTACCTCCGACACGCTGGAATGCGGTCCGAGATGGGAGCAAAATCCGAAACCTACTA 12889
|||||

QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
|||||
Db 12888 GGGCTGGCACTGGCTCGTTTGGAGGAGGAGCACAAGGCACGTGGTGTCTCAGGCTCC 12829
|||||

QY 181 GGCAGGGCTGCAGGAAAGGCTGCAGCTCGCTGAGATTGTCAAGCGGGGTCGCCAGGC 240
|||||
Db 12828 GGCAGGGCTGCAGGAAAGGCTGCAGCTCGCTGAGATTGTCAAGCGGGGTCGCCAGGC 12769
|||||

QY 241 CTGCACCACTCACCAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
|||||
Db 12768 CTACACCACTCACCAGCTTCGCTTCTTCTGCAACTGAGGACAGCTGGGTCCCAGCCTCA 12709
|||||

QY 301 CCTGACACAGGGCTAGACCCCTCAGTGGCGGCCCATGTGCTGCTGCTGCTGCTGCTG 360
|||||
Db 12708 CCGGACACGGGCTGGACCCCTCAGTCCGACGCTCCGACGCTGCTGCTGCTGCTGCTG 12649
|||||

QY 361 CTCAGCGGGGACCCCTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 420
|||||
Db 12648 CTTAGTCGGGACCCCTTGGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT 12589
|||||

QY 421 GGCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGAGAGGGGTCGAGAC 480
|||||
Db 12588 GGCTGGGTTCCATGCCAGCTGTGGTCCGAGAGGGGTCGAGAGGGGTCGAGAG 12529
|||||

QY 481 ACCCGATCGTGA 492
|||||

Db 12528 ACCCGTCTCTGA 12517
|||||

RESULT 9
MMU293897
LOCUS Mus musculus mRNA for hypothetical protein (ORF1), related to clone
Telethon(Italy_B41)_Strait02713_FL143.
ACCESSION AJ293897
VERSION AJ293897.1 GI:13397924
KEYWORDS ORF1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Ievolella,C., Zara,I. and Lanfranchi,G.
AUTHORS Full length sequencing of some human and murine muscular transcript
TITLE (Telethon Italy project B41)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 834)
AUTHORS Ievolella,C.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2000) Ievolella C., CRIBI Biotechnology
Centre-Compl. Interdepartim. Vallisneri, Universita' di Padova, via
Ugo Bassi 58/b, 35121, ITALY
FEATURES
Source
Location/Qualifiers
1. .834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="skeletal muscle"
/tissue_lib="Stratagene Unizap XR vector"
1. .93
/note="ORF1"
94. .585
/note="ORF1"

5'UTR
CDS

related to clone Telethon(Italy_B41)_Strait02713_FL143"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC34588.1"
/db_xref="GI:13397925"
/translation="MEQYRRAGSVELPASSPMPLPDDTLEMRVRDGSKIRNLLGLAL
GRLEGSTRHVFSGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPTSPD
TGLDPLTVRRHVPVAVWLLSRDPLDPSECGYQPPGAPGLGSIIPSPSCGPRRRARD
TRS"
585. .834
/note="ORF1"
834
/note="ORF1"
polyA_site
ORIGIN

Query Match	75.3%;	Score 370.4;	DB 10;	Length 834;
Best Local Similarity	84.6%;	Pred. No. 1.3e-83;		
Matches 416;	Conservative 0;	Mismatches 76;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG	60	
Db	94	ATGGAACAGTACCGGAGGCGGCTCTGTAGAGCTTCAGGCCTCATCACCAATGCCCCAG	153	
QY	61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGCGAGCAAAATTCGCAACTGCTG	120	
Db	154	CTACCTCCTGACACTCTGGAATCGAGTCCGAGATGGAGCAAAATCCGAACCTGCTA	213	
QY	121	GGGTTGGCTCTGGGTCGGTTGGAGGCGGCAGTGCCTCGGCATGTAGTGTCTCAGGTTCT	180	
Db	214	GGGCTGGCGTGGGTCGTTTGAAGGGGGAAGCACCGACACGTGGTGTCTCAGGCTCT	273	
QY	181	GGCAGGGCTGCAGGAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGCGGTCCAGGC	240	
Db	274	GGCCGGGCTGTGGAAGGCTGTGAGTGTGACAGATTGTCAAAGGCGGGTCCGGGC	333	
QY	241	CTGCACCAAGTACCAAGTACGTTTCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA	300	
Db	334	CTACACCAAGTACCAAGTCCGCTTCTGTCAAACGGAGGACAGCTGGGTCCCAACTCA	393	
QY	301	CCTGACACAGGGCTAGACCCCTCACAGTGCGCCGCGCATGTGCCTGCAGTGTGGTCTG	360	
Db	394	CCAGACACGGGACTAGACCCCTCACAGTCCGACGCCACGTCGCTGCGGTGTGGTACTG	453	
QY	361	CTCAGCCGGGACCCCTTGACCCCAATGAGTGTGTTTACCAACCCCGAGGACACCCCT	420	
Db	454	CTTAGTCGGGACCCCTTGACCCCAAGTGAATGTGGCTACCAACCCCGAGGTCACCTCCT	513	
QY	421	GGCCTGGTTCCATGCCACGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC	480	
Db	514	GGCCTGGTTCCATACCTAGTCCAGCTGTGGCCCGCAGACCCCGAAGGAGGCTCGGGAC	573	
QY	481	ACCCGATCGTGA	492	
Db	574	ACCCGGTCTTGA	585	

RESULT 10
BC047068
LOCUS
DEFINITION
Mus musculus RIKEN cdna 2810432D09 gene, mRNA (cdna clone MGC:32232
IMAGE:5009034), complete cds.
ACCESSION
BC047068
VERSION
BC047068.1 GI:28436878
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 863)
Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schéétz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 863)
Strausberg,R.
Direct Submission
Submitted (19-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 45 Row: b Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21312503.

FEATURES
Source

1. .863
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:32232 IMAGE:5009034"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .863
/gene="2810432D09Rik"
/db_xref="LocusID:69961"
/db_xref="MGI:1917211"
91. .582
/gene="2810432D09Rik"
/codon_start=1
/product="RIKEN CDNA 2810432D09"
/protein_id="AAH47068.1"
/db_xref="GI:28436879"
/db_xref="LocusID:69961"
/db_xref="MGI:1917211"
/translation="MEQYRRAGSVELPASSPMPLPDDTLEMRVRDGSKIRNLLGLAL
GRLEGSTRHVFSGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPTSPD
TGLDPLTVRRHVPVAVWLLSRDPLDPSECGYQPPGAPGLGSIIPSPSCGPRRRARD
TRS"

gene

CDS

ORIGIN

Query Match	75.3%	Score 370.4	DB 10	Length 863
Best Local Similarity	84.6%	Pred. No. 1.3e-83		
Matches 416	Conservative 0	Mismatches 76	Indels 0	Gaps 0
QY 1	ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	60		
Db				
QY 91	ATGGAACAGTACCGAGGGCGGCTCTGTAGAGCTTCCAGCCTCATCACCATGCCCCAG	150		
Db				
QY 61	CTACCTCCTGATACCTTGAGATCGGGTCCGAGATGGAGCAAAATTCGCAACCTGCTG	120		
Db				
QY 151	CTACCTCCTGACACTCTGGAATCGAGTCCGAGATGGAGCAAAATCCGAAACCTGCTA	210		
Db				
QY 121	GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTCCGATGTAGTGTCTCAGGTTCT	180		
Db				
QY 211	GGGCTGGCGCTGGTTCGTTTGAAGGGGAAGCACCCAGGACGTTGTTCTCAGGCTT	270		
Db				
QY 181	GGCAGGGCTGCAGGAAGGCTGTCACTGCGCTGAGATTGTCAAGCGCGGGTCCCAAGC	240		
Db				
QY 271	GGCCGGGCTGCTGGAAGGCTGTCACTGTGCAGAGATTGTCAAAGGGGGTTCGGGC	330		
Db				
QY 241	CTGCACCAGCTCACCAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCTCA	300		
Db				
QY 331	CTACACCAGCTCACCAGCTCCGTTCTTGCAACGGAGGACAGCTGGTCCCAACTCA	390		
Db				
QY 301	CCTGACACAGGCTAGACCCCCCTCACAGTCCGCCCATGTGCCTGCAGTGTGGTGTG	360		
Db				
QY 391	CCAGACACGGGACTAGACCCCCCTCACAGTCCGACGCGCACGTCCTGCGGTGTGGTACTG	450		
Db				
QY 361	CTCAGCCGGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCCCAGGACACCCCT	420		
Db				
QY 451	CTTAGTCGGGACCCCTTGACCCCAAGTGAATGTGGTACCAACCCCCCAGGTGCACCTCT	510		
Db				
QY 421	GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGAAGAGGGCTCAGAC	480		
Db				
QY 511	GGCCTGGGTTCCATACCTAGTCCCAGCTGTGGCCCCAGACCCCCGAGGAGGGCTCGGAC	570		
Db				
QY 481	ACCCGATCGTGA	492		
Db				
QY 571	ACCCGGTCCCTGA	582		
Db				

RESULT, T 17

AL807796/c
LOCUS AL807796 168268 bp DNA linear ROD 04-FEB-2003
DEFINITION Mouse DNA sequence from clone RP23-167I12 on chromosome 4, complete
sequence.
ACCESSION AL807796
VERSION AL807796.16 GI:28144635
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 168268)
Harrison,E.
Direct Submission
Submitted (04-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 30, 2003 this sequence version replaced gi:27652683.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk


```
RESULT 12
AX887740
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
Expressed sequence tags and encoded human proteins
Patent: EP 1033401-A 3603 06-SEP-2000;
Genset (FR)
FEATURES
source
Location/Qualifiers
1..303
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
68..>301
/note="unnamed protein product"
/codon_start=1
/protein_id="CAP01878.1"
/db_xref="GI:40046204"
/translation="MEHYRKAGSVELPAPSPMPQLPPTLEMRVRDGSKIRNLLGLAL
GRLEGG SARHVVFGSGRAAGKAVSCAEIVKRRV"
ORIGIN
Query Match 48.0%; Score 236; DB 6; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 60
|||||
Db 68 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 127
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG 120
|||||
Db 128 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG 187
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTCAAGCGGGGTCCC 236
|||||
Db 188 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTCAAGCGGGGTCCC 303
QY 181 GGCAGGGCTGCAGGAAAGCTGTCAAGTGCAGTGTCAAGCGGGGTCCC 236
|||||
Db 248 GGCAGGGCTGCAGGAAAGCTGTCAAGTGTCAAGCGGGGTCCC 303
RESULT 13
BD027350
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 303)
Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
Sequence tag and encoded human protein
Patent: JP 2001269182-A 3596 02-OCT-2001;
GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/3596
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
```

```
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 68..301.
FEATURES
source
Location/Qualifiers
1..303
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 48.0%; Score 236; DB 6; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 60
|||||
Db 68 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 127
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG 120
|||||
Db 128 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG 187
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTCAAGCGGGGTCCC 236
|||||
Db 188 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGTCAAGCGGGGTCCC 303
QY 181 GGCAGGGCTGCAGGAAAGCTGTCAAGTGCAGTGTCAAGCGGGGTCCC 236
|||||
Db 248 GGCAGGGCTGCAGGAAAGCTGTCAAGTGTCAAGCGGGGTCCC 303
RESULT 14
BD077019
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 305)
Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
5' EST of secretory protein expressed in prostate
Patent: JP 2001512013-A 266 21-AUG-2001;
GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001512013-A/266
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505291
PR 01-AUG-1997 US 08/905144
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI
LACROIX
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
blastn
CC identity 99
CC region 9..294
CC id W52125
CC est
CC blastn
CC identity 100
CC region 1..265
CC id AA024623
CC est
CC blastn
CC identity 100
CC region 1..263
```


Best Local Similarity 58.4%; Pred. No. 4.1e-11;			
Matches 149; Conservative 0; Mismatches 106; Indels 0; Gaps 0;			
QY	156	TCGGCATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTCAAGCTGCGCTGA	215
Db			
	455	TCGCCAAATAGTATTACAGGCGTGGGTTCAGAGCGTTGCAAAAGCAATCACGTGTGTAGA	514
QY	216	GATTGTCAAGCGCGGGTCCCAGGCCCTGCACCAGCTCACCAAGCTACGTTTCCTTCAGAC	275
Db			
	515	GATCATGAACGGCGCATTCATGGGCTACATCAGCTCACTAAGCTGGCCTACCGCACACT	574
QY	276	TGAGGACAGCTGGGTCCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGCG	335
Db			
	575	TCAGGATGTCTGGGAGCCGCTGGAGCCTGGGGCCGGGCTGGACAGTCTCACCGTCAGCAG	634
QY	336	CCATGTGCCCTGCAGTGTGGGTGCTGCTCAGCCGGGACCCCTGGACCCCAATGAGTGTGG	395
Db			
	635	AAATGTACCTAGTATTTGGGTTCTTCTCTCAAGAGACTCGCTTGATAAGAACCCAGCCGGG	694
QY	396	TTACCAACCCCCAGG	410
Db			
	695	TTATCAAGCTCCAGG	709

Search completed: January 31, 2005, 17:27:11
Job time : 2526 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:10:30 ; Search time 2432 Seconds
(without alignments)
7371.854 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccgaaagc.....ctcgagacacccgacgtgta 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	492	100.0	562	5	BU785328 in44b11.y
2	492	100.0	631	2	AW957840 EST369910
3	492	100.0	711	4	BG328094 602427145
4	492	100.0	791	6	CB959493 AGENCOURT
5	492	100.0	888	1	AL528391 AL528391
6	492	100.0	888	3	CR619309 full-leng
7	492	100.0	918	5	BQ218097 AGENCOURT
8	492	100.0	940	5	BQ643214 AGENCOURT
9	492	100.0	1000	5	BQ056655 AGENCOURT
10	492	100.0	1025	5	BQ056288 AGENCOURT
11	492	100.0	1074	2	BE742621 601575430
12	492	100.0	1079	5	BQ063771 AGENCOURT
13	492	100.0	1104	4	BM806218 AGENCOURT
14	491	99.8	967	5	BQ070451 AGENCOURT
15	490.4	99.7	604	6	CB110489 K-EST0151
16	490.4	99.7	964	4	BG766764 602740060
17	490.4	99.7	1171	4	BM467200 AGENCOURT
18	490	99.6	734	1	AV705461 AV705461
19	489.4	99.5	572	7	CK902681 in44b11.x
20	486.2	98.8	889	5	BU151215 AGENCOURT
21	485.6	98.7	913	5	BQ887804 AGENCOURT
22	483	98.2	717	4	BM701721 UI-E-CQ1-
23	479.4	97.4	830	4	BG766574 602739343
24	479	97.4	633	2	AW957916 EST369986

25	473	96.1	566	7	CN406927	CN406927 170006000
26	469.8	95.5	526	1	AA313780	AA313780 EST185643
27	468	95.1	492	5	BU076706	BU076706 im5le03.y
28	462	93.9	530	1	AA469996	AA469996 zu10c07.x
29	461.6	93.8	948	4	BI768478	BI768478 603053963
30	459	93.3	923	2	BE790325	BE790325 601480501
31	458.4	93.2	978	5	BU159046	BU159046 AGENCOURT
32	445	90.4	890	2	BE899574	BE899574 601682407
33	438.4	89.1	780	2	BF699406	BF699406 602125986
34	437.8	89.0	887	5	BQ224958	BQ224958 AGENCOURT
35	434.8	88.4	790	2	BE897665	BE897665 601438569
36	426	86.6	494	5	BX280954	BX280954 BX280954
C	37	410.4	83.4	5	BQ575028	BQ575028 UI-H-EZ1-
38	408.8	83.1	857	7	CN157400	CN157400 946103 MA
C	39	408.8	83.1	7	CN159345	CN159345 948407 MA
40	407.2	82.8	573	4	BG733063	BG733063 346756 MA
41	405.6	82.4	662	5	BP453103	BP453103 BP453103
C	42	405.6	82.4	5	BU689983	BU689983 UI-CF-FNO
43	405.6	82.4	745	7	CK454492	CK454492 914978 MA
44	405.6	82.4	774	5	BX671114	BX671114 BX671114
45	403.8	82.1	580	1	AA206557	AA206557 zq51d08.r

ALIGNMENTS

RESULT 1
BU785328
LOCUS in44b11.y1 HR85 islet Homo sapiens CDNA clone IMAGE:6124773 5'
DEFINITION similar to TR:Q9V9B6 Q9V9B6 CG9422 PROTEIN. ; mRNA sequence.
ACCESSION BU785328 562 bp mRNA linear EST 11-OCT-2002
VERSION BU785328.1 GI:23831081
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 562)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: in44b11.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. .562
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6124773"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: in44b11.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. .562
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6124773"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

FEATURES
source

Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinooue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

	Query Match	100.0%;	Score 492;	DB 5;	Length 562;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-115;		
	Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCACGCGCCTTCCCCAATGCCCCAG	60		
Db	9	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCACGCGCCTTCCCCAATGCCCCAG	68		
QY	61	CTACCTCTGTATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG	120		
Db	69	CTACCTCTGTATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG	128		
QY	121	GGTTGGCTCTGGTCGGTTGGAGGGCGGCGAGTGTCCGCGATGTAGTGTCTCAGGTTCT	180		
Db	129	GGTTGGCTCTGGTCGGTTGGAGGGCGGCGAGTGTCCGCGATGTAGTGTCTCAGGTTCT	188		
QY	181	GGCAGGGCTGCAGGAAAGGCTGTCACTGCGCTGAGATTGTCAAGCGGGCGGTCCCAGGC	240		
Db	189	GGCAGGGCTGCAGGAAAGGCTGTCACTGCGCTGAGATTGTCAAGCGGGCGGTCCCAGGC	248		
QY	241	CTGCACCAGCTCACAAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA	300		
Db	249	CTGCACCAGCTCACAAAGCTACGTTTCTTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA	308		
QY	301	CCTGACACAGGCTAGACCCCTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG	360		
Db	309	CCTGACACAGGCTAGACCCCTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG	368		
QY	361	CTCAGCCGGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCCAGGACACCCCTT	420		
Db	369	CTCAGCCGGGACCCCTTGACCCCAATGAGTGTGGTTACCAACCCCCAGGACACCCCTT	428		
QY	421	GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC	480		
Db	429	GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC	488		
QY	481	ACCCGATCGTGA	492		
Db	489	ACCCGATCGTGA	500		

RESULT 2
AW957840
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW957840 631 bp mRNA linear EST 01-JUN-2000
EST369910 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
AW957840
AW957840.1 GI:8147523
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 631)
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208

Email: johnq@tigr.org
 Plate: 111
 Seq primer: Reverse.

FEATURES Location/Qualifiers
source 1. .631

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE reses
/note="Vector: pBlueses

```

ORIGIN

	Query Match	100.0%;	Score 492;	DB 2;	Length 631;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-115;		
	Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	60		
DB	52	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	111		
QY	61	CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGAAACCTGCTG	120		
DB	112	CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGAAACCTGCTG	171		
QY	121	GGTTGGCTCTGGGTTCGGTTGGAGGGCGGAGTGCTCGGCATGTAGTGTCTCAGGTCT	180		
DB	172	GGTTGGCTCTGGGTTCGGTTGGAGGGCGGAGTGCTCGGCATGTAGTGTCTCAGGTCT	231		
QY	181	GGCAGGGCTGCAGGAAAGGCTGTCTAGCTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC	240		
DB	232	GGCAGGGCTGCAGGAAAGGCTGTCTAGCTGCGCTGAGATTGTCAAGCGCGGGTCCCAGGC	291		
QY	241	CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA	300		
DB	292	CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA	351		
QY	301	CCTGACACAGGGCTAGACCCCCCTACAGTGCGCCGCCATGTGCCCTGCAGTGGGTGCTG	360		
DB	352	CCTGACACAGGGCTAGACCCCCCTACAGTGCGCCGCCATGTGCCCTGCAGTGGGTGCTG	411		
QY	361	CTCAGCCGGACCCCCCTGGACCCCAATGAGTGTGGTTACCAACCCCCCAGGAGCACCCCCCT	420		
DB	412	CTCAGCCGGACCCCCCTGGACCCCAATGAGTGTGGTTACCAACCCCCCAGGAGCACCCCCCT	471		
QY	421	GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCCCTCGTTCCCGAAGAGGGCTCGAGAC	480		
DB	472	GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCCCTCGTTCCCGAAGAGGGCTCGAGAC	531		
QY	481	ACCCGATCGTGA	492		
DB	532	ACCCGATCGTGA	543		

RESULT 3	REFERENCE
BG328094	AUTHORS
LOCUS	TITLE
DEFINITION	JOURNAL
ACCESSION	COMMENT
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Ho

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1233 row: i column: 12
High quality sequence stop: 706.
Location/Qualifiers

FEATURES
source

1. .711
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4546643"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

ORIGIN

Query Match 100.0%; Score 492; DB 4; Length 711;
Best Local Similarity 100.0%; Pred. No. 2.7e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 113
QY 61 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGAGTGGTGGCATGTAGTTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGAGTGGTGGCATGTAGTTCTCAGGTTCT 233
QY 181 GGCAGGGCTGCAGGAAAGCTGTACGCTGGCTGAGATTGTCAAGCGGGGTCCCAGGC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 GGCAGGGCTGCAGGAAAGGCTGTACGCTGGCTGAGATTGTCAAGCGGGGTCCCAGGC 293
QY 241 CTGACCAGCTACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
294 CTGACCAGCTACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 353
QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGGCCCATGTGCCTGCAGTGTGGGTGCTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
354 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGGCCCATGTGCCTGCAGTGTGGGTGCTG 413
QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCAAGGAGCACCCCT 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
414 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCAAGGAGCACCCCT 473
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
474 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGCTCGAGAC 533
QY 481 ACCCGATCGTGA 492
Db ||||||||||||
534 ACCCGATCGTGA 545

RESULT 4
CB959493
LOCUS CB959493 791 bp mRNA linear EST 29-APR-2003
DEFINITION AGENCOURT_13887473 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30341789 5', mRNA sequence.
ACCESSION CB959493

VERSION

KEYWORDS CB959493.1 GI:30215609

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Homo sapiens

AUTHORS 1 (bases 1 to 791)

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help

and advice from Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM373 row: n column: 06

High quality sequence stop: 631.

FEATURES

Location/Qualifiers

1. .791

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30341789"

/tissue_type="Human Placenta"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_147"

/note="Organ: placenta; Vector: pBluescriptR; Site_1:

ali-XhoI; Site_2: BamH; Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: This is

a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 492; DB 6; Length 791;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
95 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAGCGCCTTCCCCAATGCCCCAG 154
QY 61 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
155 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 214
QY 121 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCGAGTGTGCGCATGTAGTTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
215 GGGTTGGCTCTGGGTTCGGTTGGAGGGCGGCGAGTGTGCGCATGTAGTTCTCAGGTTCT 274
QY 181 GGCAGGGCTGCAGGAAAGGCTGTACGCTGCGCTGAGATTGTCAAGCGGGGTCCCAGGC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
275 GGCAGGGCTGCAGGAAAGGCTGTACGCTGCGCTGAGATTGTCAAGCGGGGTCCCAGGC 334
QY 241 CTGCACCAGCTACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
335 CTGCACCAGCTACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 394
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCGCTGCGGTGCTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
395 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCGCTGCGGTGCTG 454
QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
455 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 514

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGGAAGAAGGGCTCGAGAC 480
Db 515 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGGAAGAAGGGCTCGAGAC 574

QY 481 ACCCGATCGTGA 492
Db 575 ACCCGATCGTGA 586

RESULT 5
AL528391
LOCUS
DEFINITION
AL528391 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC029YK21 5-PRIME, mRNA sequence.

ACCESSION
AL528391
VERSION
AL528391.3 GI:45703459
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:31066241.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5101.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC029AF11QPl&c=5101.r.

FEATURES
source
1. .888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC029YK21"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 100.0%; Score 492; DB 1; Length 888;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 87 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 146

QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 147 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 206

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTTCTCAGGTTCT 180
Db 207 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTGCTCGGCATGTAGTTCTCAGGTTCT 266

QY 181 GGCAGGGCTGCAGGAAAGGCTGTACGTCGGCTGAGATTGTCAAGCGGGCGGGTCCCAGGC 240
Db 267 GGCAGGGCTGCAGGAAAGGCTGTACGTCGGCTGAGATTGTCAAGCGGGCGGGTCCCAGGC 326

QY 241 CTGCACCAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 327 CTGCACCAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 386

QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGCGCCCATGTGCTGCAGTGTGGGTGCTG 360
Db 387 CCTGACACAGGGCTAGACCCCTCACAGTGCGCCCATGTGCTGCAGTGTGGGTGCTG 446

QY 361 CTCAGCCGGGACCCCTTGAGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCT 420
Db 447 CTCAGCCGGGACCCCTTGAGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCT 506

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGGAAGAAGGGCTCGAGAC 480
Db 507 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCTCCGGAAGAAGGGCTCGAGAC 566

QY 481 ACCCGATCGTGA 492
Db 567 ACCCGATCGTGA 578

RESULT 6
CR619309
LOCUS
DEFINITION
CR619309 888 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DC029YK21 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).

ACCESSION
CR619309
VERSION
CR619309.1 GI:50500116
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished

AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
Full-length cDNA libraries and normalization

JOURNAL
Unpublished

REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue

REFERENCE
2 (bases 1 to 888)
Genoscope.
Direct Submission

AUTHORS
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
Location/Qualifiers
1. .888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC029YK21"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 100.0%; Score 492; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 87 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 146

QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 147 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 206


```
QY 121 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 207 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCAGTCTCGGCATGTAGTGTCTCAGGTTCT 266
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTCGCGCTGAGATTGTCAAGCGGCGGGTCCCAGGC 240
Db 267 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTCGCGCTGAGATTGTCAAGCGGCGGGTCCCAGGC 326
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 327 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 386
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCGCCATGTCCCTGCAAGTGGGTGCTG 360
Db 387 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCGCCATGTCCCTGCAAGTGGGTGCTG 446
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 447 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 506
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 480
Db 507 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 566
QY 481 ACCCGATCGTGA 492
Db 567 ACCCGATCGTGA 578
```

```
RESULT 7
BQ218097
LOCUS
DEFINITION BQ218097 7510498 NIH_MGC_70 Homo sapiens cDNA clone EST 02-MAY-2002
5', mRNA sequence.
```

```
ACCESSION BQ218097
VERSION BQ218097.1 GI:20399497
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
```

```
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13219 row: d column: 04
High quality sequence stop: 539.
```

```
FEATURES
source
1..918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6018387"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
```

ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 918;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;

```
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGGCGCTTCCCCAATGCCCCAG 60
Db 58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGGCGCTTCCCCAATGCCCCAG 117
QY 61 CTACCTCCTGATACCTTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCCGAACCTGCTG 120
Db 118 CTACCTCCTGATACCTTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCCGAACCTGCTG 177
QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 178 GGGTTGGCTCTGGGTGCGTTGGAGGGCGGCAGTCTCGGCATGTAGTGTCTCAGGTTCT 237
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTCGGCTGAGATTGTCAAGCGGCGGGTCCCAGGC 240
Db 238 GGCAGGGCTGCAGGAAAGGCTGTCAAGCTCGGCTGAGATTGTCAAGCGGCGGGTCCCAGGC 297
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 298 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 357
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCGCCATGTCCCTGCAAGTGGGTGCTG 360
Db 358 CCTGACACAGGGCTAGACCCCTCACAGTGCAGCGCCATGTCCCTGCAAGTGGGTGCTG 417
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 418 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 477
QY 421 GGCCTGGGTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 480
Db 478 GGCCTGGGTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 537
QY 481 ACCCGATCGTGA 492
Db 538 ACCCGATCGTGA 549
```

```
RESULT 8
BQ643214
LOCUS
```

```
DEFINITION BQ643214 8485520 NIH_MGC_99 Homo sapiens cDNA clone EST 15-JUL-2002
5', mRNA sequence.
```

```
ACCESSION BQ643214
VERSION BQ643214.1 GI:21767386
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM2527 row: j column: 21
High quality sequence stop: 615.

FEATURES
source

1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6305396"
/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 940;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 113
QY 61 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGCGGTCCCAGGC 293
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
294 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 353
QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
354 CCTGACACAGGGCTAGACCCCTTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG 413
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
414 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 473
QY 421 GGCCTGGTTCCATGCCCCAGCTCCAGCTGTGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
474 GGCCTGGTTCCATGCCCCAGCTCCAGCTGTGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 533
QY 481 ACCCGATCGTGA 492
Db ||||||||||||
534 ACCCGATCGTGA 545

RESULT 9
BQ056655

LOCUS
DEFINITION
AGENCOURT_6940426 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5809057
5', mRNA sequence.

ACCESSION
BQ056655
VERSION
BQ056655.1 GI:19815982
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1000)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2054 row: b column: 02
High quality sequence stop: 652.
Location/Qualifiers
1..1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5809057"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

FEATURES
source

ORIGIN

Query Match 100.0%; Score 492; DB 5; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
50 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 109
QY 61 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
110 CTACCTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 169
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
170 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGTGCTGGCATGTAGTGTCTCAGGTTCT 229
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTGCGTGCAGTGTGAGTGTCAAGCGGCGGTCCCAGGC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
230 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTGCGTGCAGTGTGAGTGTCAAGCGGCGGTCCCAGGC 289
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
290 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 349
QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
350 CCTGACACAGGGCTAGACCCCTTCACAGTGCGCCGCCATGTGCCTGCAGTGTGGTGCTG 409
QY 361 CTGAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
410 CTGAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCT 469
QY 421 GGCCTGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
470 GGCCTGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC 529
QY 481 ACCCGATCGTGA 492
Db ||||||||||||
530 ACCCGATCGTGA 541

RESULT 10
BQ056288
LOCUS
DEFINITION
AGENCOURT_6773378 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808601

BQ056288 1025 bp mRNA linear EST 29-MAR-2002
AGENCOURT_6773378 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808601

Db 213 GGCAGGGCTGCAGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 272
QY 241 CTGCACCAAGCTACAAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db 273 CTGCACCAAGCTACAAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 332
QY 301 CCTGACACAGGGTAGACCCCTCAGAGTGGCGGCCCATGTGCCTGAGTGTGGGTGCTG 360
Db 333 CCTGACACAGGGTAGACCCCTCAGAGTGGCGGCCCATGTGCCTGAGTGTGGGTGCTG 392
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 393 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 452
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 453 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 512
QY 481 ACCCGATCGTGA 492
Db 513 ACCCGATCGTGA 524

RESULT 12
BQ063771
LOCUS
DEFINITION
AGENCOURT_6831996 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925418
5', mRNA sequence.
BQ063771
VERSION
BQ063771.1 GI:19891815
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1079)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM2097 row: b column: 11
High quality sequence stop: 520.
Location/Qualifiers
1. .1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925418"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

FEATURES
source
1. .1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925418"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 100.0%; Score 492; DB 5; Length 1079;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;

Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 54 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 113
QY 61 CTACCTCCTGATACCCCTTGAGATGGGGTCCGAGATGGCAAAAATTCGCAACCTGCTG 120
Db 114 CTACCTCCTGATACCCCTTGAGATGGGGTCCGAGATGGCAAAAATTCGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 174 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db 234 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 293
QY 241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGTGGGTCCCAGCCTCA 300
Db 294 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGTGGGTCCCAGCCTCA 353
QY 301 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCCATGTGCCTGCGGTGCTG 360
Db 354 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCCATGTGCCTGCGGTGCTG 413
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 414 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 473
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 474 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 533
QY 481 ACCCGATCGTGA 492
Db 534 ACCCGATCGTGA 545

RESULT 13
BQ06218
LOCUS
DEFINITION
AGENCOURT_6554071 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555733
5', mRNA sequence.
BQ06218
VERSION
BQ06218.1 GI:19123029
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1104)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12276 row: f column: 22
High quality sequence stop: 630.
Location/Qualifiers
1. .1104
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5555733"
/tissue_type="leiomyosarcoma"


```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally.. Primer: Oligo dT.
Average insert size 2.1 kb. "

ORIGIN
Query Match      100.0%; Score 492; DB 4; Length 1104;
Best Local Similarity 100.0%; Pred. No. 2.9e-115;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 163 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 222
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 120
Db 223 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 282
QY 121 GGGTTGGCTCTGGTTCGGTTCGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 283 GGGTTGGCTCTGGTTCGGTTCGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 342
QY 181 GGCAGGGCTGCAGGAAAGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGGCGGTCCAGGC 240
Db 343 GGCAGGGCTGCAGGAAAGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGGCGGTCCAGGC 402
QY 241 CTGCACCAAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 300
Db 403 CTGCACCAAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCA 462
QY 301 CCTGACACAGGGCTAGACCCCTCAGCTGCGCGCGCATGTGCCTGCAGTGTGGGTGCTG 360
Db 463 CCTGACACAGGGCTAGACCCCTCAGCTGCGCGCGCATGTGCCTGCAGTGTGGGTGCTG 522
QY 361 CTCAGCCGGGACCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACACCCCTT 420
Db 523 CTCAGCCGGGACCCCTGAGCCCAATGAGTGTGGTTCACCAACCCCGAGGACACCCCTT 582
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 583 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 642
QY 481 ACCCGATCGTGA 492
Db 643 ACCCGATCGTGA 654

RESULT 14
BQ070451
LOCUS
DEFINITION
AGENCOURT_6839400 NIH_MGC_47 Homo sapiens cdna clone IMAGE:5922837
5', mRNA sequence.
ACCESSION
BQ070451
VERSION
BQ070451.1 GI:19899497
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 967)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```

```

Plate: LLCM2090 row: f column: 22
High quality sequence stop: 678.
FEATURES
source
1..967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5922837"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      99.8%; Score 491; DB 5; Length 967;
Best Local Similarity 99.8%; Pred. No. 5.1e-115;
Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 254 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 313
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 120
Db 314 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 373
QY 121 GGGTTGGCTCTGGTTCGGTTCGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 180
Db 374 GGGTTGGCTCTGGTTCGGTTCGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT 433
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGCCTGAGATTGTCAAGCGGCGGTCCCAGGC 240
Db 434 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGCCTGAGATTGTCAAGCGGCGGTCCCAGGC 493
QY 241 CTGCACCAAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 494 CTGCACCAAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 553
QY 301 CCTGACACAGGGCTAGACCCCTCAGAGTGGCCCGCCATGTGCCTGCAGTGTGGGTGCTG 360
Db 554 CCTGACACAGGGCTAGACCCCTCAGAGTGGCCCGCCATGTGCCTGCAGTGTGGGTGCTG 613
QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCTT 420
Db 614 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCTT 673
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGCGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 674 GGCCTGNGTTCCATGCCAGCTCCAGCTGCGCCCTCGTTCCCGAAGAGGGCTCGAGAC 733
QY 481 ACCCGATCGTGA 492
Db 734 ACCCGATCGTGA 745

RESULT 15
CB110489
LOCUS
DEFINITION
K-EST0151955 L5HLK1 Homo sapiens cdna clone L5HLK1-5-D07 5', mRNA
sequence.
ACCESSION
CB110489
VERSION
CB110489.1 GI:27936296
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 604)

REFERENCE
AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE
JOURNAL
COMMENT

21C Frontier Korean EST Project 2001
Unpublished (2002)

Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 5 row: D column: 07

High quality sequence stop: 604.

FEATURES

source

1..604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L5HLK1-5-D07"
/sex="M"
/cell_line="HLK-1"
/lab_host="Top10F"
/clone_lib="L5HLK1"
/note="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 99.7%; Score 490.4; DB 6; Length 604;
Best Local Similarity 99.8%; Pred. No. 6.9e-115;
Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	60
Db	69	ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	128
QY	61	CTACCTCCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG	120
Db	129	CTACCTCCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG	188
QY	121	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT	180
Db	189	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT	248
QY	181	GGCAGGGCTGCAGAAAGGCTGTCAGCTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC	240
Db	249	GGCAGGGCTGCAGAAAGGCTGTCAGCTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC	308
QY	241	CTGCACCAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGTGGGTCCCAGCCTCA	300
Db	309	CTGCACCAGCTACCAAGCTACGTTTCTTCAGACTGAGGACAGTGGGTCCCAGCCTCA	368
QY	301	CCTGACACAGGGCTAGACCCCTCACAGTGGCCGCCCATGTGCCTGCAGTGTGGTGCTG	360
Db	369	CCAGACACAGGGCTAGACCCCTCACAGTGGCCGCCCATGTGCCTGCAGTGTGGTGCTG	428

QY	361	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	420
Db	429	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	488
QY	421	GGCCTGGGTTCCATGCCAGCTCCAGCTCCAGCTGTGGCCCTCGTTCCGAAAGGGCTCGAGAC	480
Db	489	GGCCTGGGTTCCATGCCAGCTCCAGCTCCAGCTGTGGCCCTCGTTCCGAAAGGGCTCGAGAC	548
QY	481	ACCCGATCGTGA	492
Db	549	ACCCGATCGTGA	560

Search completed: January 31, 2005, 18:07:49
Job time : 2438 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 17:27:15 ; Search time 423 Seconds
(without alignments)
6683.161 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgtagtga 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues 8600550
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq: *
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	492	100.0	894	9	US-09-731-872-218	Sequence 218, App
2	492	100.0	894	10	US-09-876-997-218	Sequence 218, App
3	484	98.4	884	11	US-09-978-360A-311	Sequence 311, App
4	373.8	76.0	469	10	US-09-918-995-21099	Sequence 21099, A
5	368.4	74.9	1168	9	US-09-981-353-1	Sequence 1, Appli
6	368.4	74.9	1168	15	US-10-084-817-60	Sequence 60, Appl
7	223.8	45.5	560	16	US-10-152-319A-920	Sequence 920, App
8	197.8	40.2	299	18	US-10-425-115-52678	Sequence 52678, A
9	74.6	15.2	1941	15	US-10-037-270-5	Sequence 5, Appli
10	74.6	15.2	1941	15	US-10-117-722-5	Sequence 5, Appli
11	68.2	13.9	412	9	US-09-983-965-2039	Sequence 2039, Ap
12	40.4	8.2	929	18	US-10-653-047-7620	Sequence 7620, Ap

C 13	39.4	8.0	32189	9	US-09-764-877-3291	Sequence 3291, Ap
C 14	39.4	8.0	32189	16	US-10-242-515-3291	Sequence 3291, Ap
C 15	38.8	7.9	972	15	US-10-156-761-4789	Sequence 4789, Ap
16	38.8	7.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
C 17	37.8	7.7	594	14	US-10-123-155-10	Sequence 10, Appl
C 18	37.8	7.7	594	15	US-10-146-731-10	Sequence 10, Appl
C 19	37.8	7.7	594	15	US-10-140-472-10	Sequence 10, Appl
C 20	37.8	7.7	594	15	US-10-141-761-10	Sequence 10, Appl
C 21	37.8	7.7	594	15	US-10-142-885-10	Sequence 10, Appl
C 22	37.8	7.7	594	15	US-10-158-790-10	Sequence 10, Appl
C 23	37.8	7.7	594	15	US-10-137-871-10	Sequence 10, Appl
C 24	37.8	7.7	594	15	US-10-140-923-10	Sequence 10, Appl
C 25	37.8	7.7	594	15	US-10-141-756-10	Sequence 10, Appl
C 26	37.8	7.7	594	15	US-10-141-759-10	Sequence 10, Appl
C 27	37.8	7.7	594	15	US-10-140-805-10	Sequence 10, Appl
C 28	37.8	7.7	594	15	US-10-140-864-10	Sequence 10, Appl
C 29	37.8	7.7	594	16	US-10-142-426-10	Sequence 10, Appl
C 30	37.4	7.6	645	15	US-10-156-761-5727	Sequence 5727, Ap
C 31	37.4	7.6	9025608	15	US-10-156-761-1	Sequence 1, Appli
32	37.2	7.6	1218	16	US-10-425-114-21994	Sequence 21994, A
33	37.2	7.6	1520	18	US-10-425-115-16813	Sequence 16813, A
C 34	37.2	7.6	46819	14	US-10-114-170-72	Sequence 72, Appl
35	36.6	7.4	1593	15	US-10-156-761-1221	Sequence 1221, Ap
36	36.2	7.4	500	10	US-09-770-961-441	Sequence 441, App
37	36	7.3	1363	18	US-10-842-005A-2	Sequence 2, Appli
C 38	36	7.3	11100	15	US-10-329-079-5	Sequence 5, Appli
C 39	36	7.3	13315	15	US-10-329-079-1	Sequence 1, Appli
40	35.6	7.2	831	15	US-10-156-761-4034	Sequence 4034, Ap
41	35.6	7.2	22306	14	US-10-114-170-251	Sequence 251, App
42	35.4	7.2	493	9	US-09-783-590-3619	Sequence 3619, Ap
43	35.4	7.2	1993	18	US-10-425-115-79876	Sequence 79876, A
44	35.4	7.2	2662	9	US-09-925-301-2	Sequence 2, Appli
45	35.4	7.2	4257	9	US-09-825-288A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-731-872-218
; Sequence 218, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731.872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 218
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..566
US-09-731-872-218

Query Match 100.0%; Score 492; DB 9; Length 894;
Best Local Similarity 100.0%; Pred. No. 3.4e-140;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 137

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGCGCAGCAAAATTCGCAACCTGCTG 120
Db 138 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGCGCAGCAAAATTCGCAACCTGCTG 197
QY 121 GGGTTGGCTCTGGTCTGGAGGGCGGCGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 180
Db 198 GGGTTGGCTCTGGTCTGGAGGGCGGCGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 257
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 240
Db 258 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 317
QY 241 CTGCACCAAGCTACGTTTCTTCCAGTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 300
Db 318 CTGCACCAAGCTACGTTTCTTCCAGTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 377
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGTGCGCGCATGTGCGTGTGGGTGCTG 360
Db 378 CCTGACACAGGGCTAGACCCCTTACAGTGTGCGCGCATGTGCGTGTGGGTGCTG 437
QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGCGCGCATGTGCGTGTGGGTGCTG 420
Db 438 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGCGCGCATGTGCGTGTGGGTGCTG 497
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGCGCGCATGTGCGTGTGGGTGCTG 480
Db 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGCGCGCATGTGCGTGTGGGTGCTG 557
QY 481 ACCCGATCGTGA 492
Db 558 ACCCGATCGTGA 569

RESULT 2
US-09-876-997-218
; Sequence 218, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 218
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..566
US-09-876-997-218

Query Match 100.0%; Score 492; DB 10; Length 894;
Best Local Similarity 100.0%; Pred. No. 3.4e-140;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 78 ATGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 138 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197

QY 121 GGGTTGGCTCTGGTCTGGAGGGCGGCGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 180
Db 198 GGGTTGGCTCTGGTCTGGAGGGCGGCGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 257
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 240
Db 258 GGCAGGGCTGCAGGAAAGGCTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 317
QY 241 CTGCACCAAGCTACGTTTCTTCCAGTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 300
Db 318 CTGCACCAAGCTACGTTTCTTCCAGTGTGAGTGTCCGCGCATGTAGTGTCTCAGGTTCT 377
QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGTGCGCGCATGTGCGTGTGGGTGCTG 360
Db 378 CCTGACACAGGGCTAGACCCCTTACAGTGTGCGCGCATGTGCGTGTGGGTGCTG 437
QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGCGCGCATGTGCGTGTGGGTGCTG 420
Db 438 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGCGCGCATGTGCGTGTGGGTGCTG 497
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGCGCGCATGTGCGTGTGGGTGCTG 480
Db 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGCGCGCATGTGCGTGTGGGTGCTG 557
QY 481 ACCCGATCGTGA 492
Db 558 ACCCGATCGTGA 569

RESULT 3
US-09-978-360A-311
; Sequence 311, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 311
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```

; LOCATION: 69..557
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 69..224
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.69999980926514
; OTHER INFORMATION: seq LGLALGRLEGSA/RH
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 849..854
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 870..883
US-09-978-360A-311

Query Match      98.4%; Score 484; DB 11; Length 884;
Best Local Similarity 97.8%; Pred. No. 9.3e-138;
Matches 481; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY      1  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db      69  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 128

QY      61  CTACCTCTGTATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db      129  CTACCTCTGTATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 188

QY      121  GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db      189  GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 248

QY      181  GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGGGTCCCAGGC 240
Db      249  GGCAGGGCTGCAGGAAAGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGGGTCCCAGGC 308

QY      241  CTGCACCAGCTCACCAGCTACGTTTCCTTCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db      309  CTGCACCAGCTCACCAGCTACKTTTCCTTCAAACTGAGGACAGCTGGTCCCAGCCTCA 368

QY      301  CCTGACACAGGGCTAGACCCCTCACAGTGCGCCGCCATGTGCTCCTGCAGTGTGGGTGCTG 360
Db      369  CCTGACACAGGGCTARACCCCTCACAGTGCGCCGCCATGTGCTCCTGCAGTGTGGGTGCTG 428

QY      361  CTGAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCCAGGAGCACCCCT 420
Db      429  CTCASCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCCAGGAGCACCCCT 488

QY      421  GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db      489  GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCRAAAAGGGCTCRARAC 548

QY      481  ACCCGATCGTGA 492
Db      549  ACCCGATCGTGA 560

```

```

; NAME/ABI: MISC leature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 184081.24
US-09-981-353-1

Query Match      74.9%; Score 368.4; DB 9; Length 1168;
Best Local Similarity 99.7%; Pred. No. 1.9e-102;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 123 GTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTTCACAGTTCTGG 182
      |||
Db 871 GCTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTTCACAGTTCTGG 812
      |||

QY 183 CAGGGCTGCAGGAAAGGCTGTCAAGCGGGCGGGTCCAGGCCT 242
      |||
Db 811 CAGGGCTGCAGGAAAGGCTGTCAAGCGGGCGGGTCCAGGCCT 752
      |||

QY 243 GCACCAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGGCCTCACC 302
      |||

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-21099

Query Match      76.0%; Score 373.8; DB 10; Length 469;
Best Local Similarity 99.2%; Pred. No. 4e-104;
Matches 386; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      1  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG 60
      |||
Db      81  ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG 140

QY      61  CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
      |||
Db     141  CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 200

QY     121  GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAG-GTTC 179
      |||
Db     201  GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAGAGTTC 260

QY     180  TGGCAGGGCTGCAGGAAAGCTGTTCAGCTCGCTGAGATTGTCAAGCGGCGGGTCCCAGG 239
      |||
Db     261  TGGCAGGGCTGCAGGAAAGCTGTTCAGCTCGCTGAGATTGTCAAGCGGCGGGTCCCAGG 320

QY     240  CCTGCACCAGCTCACCAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGGCTC 299
      |||
Db     321  CCTGCACCAGCTCACCAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGGCTC 380

QY     300  ACCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCCCTGCAGTGTGGGTGCT 359
      |||
Db     381  ACCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCCCTGCAGTGTGGGTGCT 440

QY     360  GCTCAGCCGGGACCCCTGGACCCCAATG 388
      |||
Db     441  GCTCATCCGGGACCCCTGGACCCCAATG 469

RESULT 5
US-09-981-353-1/c
; Sequence 1, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 184081.24
US-09-981-353-1

```


Db 751 GCACGAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACC 692
QY 303 TGACACAGGGGTAGACCCCTCACAAGTGGCGGCCCATGTGCCTGCAAGTGTGGGTGCTGCT 362
Db 691 TGACACAGGGGTAGACCCCTCACAAGTGGCGGCCCATGTGCCTGCAAGTGTGGGTGCTGCT 632
QY 363 CAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCAAGGAGCACCCCTGG 422
Db 631 CAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCAAGGAGCACCCCTGG 572
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGACAC 482
Db 571 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGACAC 512
QY 483 CCGATCGTGA 492
Db 511 CCGATCGTGA 502

RESULT 6
US-10-084-817-60/c
; Sequence 60, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 1168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 184081.24
US-10-084-817-60

Query Match 74.9%; Score 368.4; DB 15; Length 1168;
Best Local Similarity 99.7%; Pred. No. 1.9e-102;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 123 GTTGGCTCTGGTCGGTTGGAGGGCGGCAAGTGTGATGTGTTCTCAGGTTCTGG 182
Db 871 GCTGGCTCTGGTCGGTTGGAGGGCGGCAAGTGTGATGTGTTCTCAGGTTCTGG 812
QY 183 CAGGGCTGCAGGAAGGCTGTCAGCTGGCTGAGATTGTCAAGCGGCGGGTCCCAGGCCT 242
Db 811 CAGGGCTGCAGGAAGGCTGTCAGCTGGCTGAGATTGTCAAGCGGCGGGTCCCAGGCCT 752
QY 243 GCACGAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACC 302
Db 751 GCACGAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACC 692
QY 303 TGACACAGGGGTAGACCCCTCACAAGTGGCGGCCCATGTGCCTGCAAGTGTGGGTGCTGCT 362
Db 691 TGACACAGGGGTAGACCCCTCACAAGTGGCGGCCCATGTGCCTGCAAGTGTGGGTGCTGCT 632
QY 363 CAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCAAGGAGCACCCCTGG 422
Db 631 CAGCCGGGACCCCTGGACCCCAATAGTGTGGTTACCAACCCCAAGGAGCACCCCTGG 572
QY 423 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGACAC 482
Db 571 CCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGACAC 512

QY 483 CCGATCGTGA 492
Db 511 CCGATCGTGA 502
RESULT 7
US-10-152-319A-920/c
; Sequence 920, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 920
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. A1104296
US-10-152-319A-920

Query Match 45.5%; Score 223.8; DB 16; Length 560;
Best Local Similarity 85.6%; Pred. No. 2.7e-58;
Matches 249; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 202 GTCAGCTGCGCTGAGATTGTCAAGCGGGGGTCCCAGGGCTGCACCAGCTCACCAGCTA 261
Db 560 GTCAGCTGTGAGAGATTGTCAACGGCGGTTTCCGGGCCCTACCCAGCTCACCAGCTT 501
QY 262 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCTCACCTGACACAGGGCTAGACCCC 321
Db 500 CGTTTCCTTCAGAACTGAGGACAGCTGGGTCCCAGCTCACCTGACACAGGGCTTGGACCCC 441
QY 322 CTCACAGTGGCGGCCCATGTGCCTGAGTGTGGGTGCTGCTAGCGGGGACCCCTGGAC 381
Db 440 CTCACAGTCCGACGCCATGTGCCTGAGTGTGGGTGCTGCTAGTGGGACCCCTTGGAC 381
QY 382 CCCAATGAGTGTGGTTACCAACCCCAAGGAGCACCCCTGGCTGGGTTCATGCCCAGC 441
Db 380 CCCAGTGAATGTGGTACCAACCCCAAGGTGCACCTCCTGGCTGGGTCCATACCTAGT 321
QY 442 TCCAGCTGTGGCCCTCGTTTCCCGAAGAAGGGCTCGAGACACCCGATCGTGA 492

Db 320 TCACGCTGTGGTCCAGACCCCGAAGGAGGGCTCGGAGACCCCGGTCCTGA 270

RESULT 8

US-10-425-115-52678

; Sequence 52678, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 52678

; LENGTH: 299

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_148039C.1

US-10-425-115-52678

Query Match 40.2%; Score 197.8; DB 18; Length 299;

Best Local Similarity 94.5%; Pred. No. 2.2e-50;

Matches 205; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCAG 60

Db 83 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCACTGCCCCAG 142

QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCCGCAACCTGCTG 120

Db 143 CTACCTCCTGATATCCTTGAGATGCGGGTCCGATGGCAGCAAAATTCCGCAACCTGCTG 202

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGCTCGGCATGTAGTGTTCAGGTTCT 180

Db 203 GGGTTGGCTCAGGGTCGGTTAGAGGGCGGAGTGCTGTGCAAGTAGTGCTCTCAGGTTCT 262

QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCTGAGA 217

Db 263 GGCAGGGCTGCAGGAAATGCTGTGAGTTCGCTGAGA 299

RESULT 9

US-10-037-270-5

; Sequence 5, Application US/10037270

; Publication No. US20030104529A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: Tillinghast, John

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/10/037,270

; CURRENT FILING DATE: 2002-01-04

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1104

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 5

; LENGTH: 1941

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (138)..(773)

US-10-037-270-5

Query Match 15.2%; Score 74.6; DB 15; Length 1941;

Best Local Similarity 57.5%; Pred. No. 1.1e-12;

Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 CCCAGCGCCTTCCCCAATGCCCAAGTACCTCTGATACCCCTTGAGATGCGGTCGAGGA 95

Db 239 CCCGGGCTCCGGCCGTTTCGAGACCTGGCGCGGCGCGGTGCACATGCGGTCGAAGGA 298

QY 96 TGGCAGCAAAATTGCAACCTGCTGGGGTTGGTCTTGGGTCGGTTGGAGGGCGGAGTGC 155

Db 299 AGGCAGCAAGATCCGGAACCTGATGGCCTTCGCCACCGCCAGCATGGCGCAGCCACCC 358

QY 156 TCGGCATGTAGTGTCTCAGGTTCTGGCAGGCTCCAGGCTGCAGCAAGGCTGTGAGTGGCGCTGA 215

Db 359 GCGCGCCATCGTCTTCAGCGGCTGCGGCGGCGGCCACCAACCCGTACGTCGCGCGGA 418

QY 216 GATTGTCAAGCGCGGGTCCAGGCTCCAGGCTGCAGCAAGGCTGCAGCAAGGCTGTGAGTGGCGCTACC 268

Db 419 GATCCTCAAGCGCGCCTGGCGGCTGCAGGCTGCAGGCTGCAGGCTGCAGGCTGCAGGCTACC 471

RESULT 10

US-10-117-722-5

; Sequence 5, Application US/10117722

; Publication No. US20030219744A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/10/117,722

; CURRENT FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 09/620,312

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1104

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 5

; LENGTH: 1941

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (138)..(773)

US-10-117-722-5

Query Match 15.2%; Score 74.6; DB 15; Length 1941;

Best Local Similarity 57.5%; Pred. No. 1.1e-12;

Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Db 25379 GGCTCCTGGGTCCCGGGCTTCTGGGTCCCCCTTCTCAGGTGCGCCTGGCCACCTTCTGG 25320
QY 431 CCATGCCAGCTCCAGC 447
Db 25319 CCTTCCCTGCACCAAC 25303

RESULT 14
US-10-242-515-3291/c
; Sequence 3291, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3291
; LENGTH: 32189
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-242-515-3291

Query Match 8.0%; Score 39.4; DB 16; Length 32189;
Best Local Similarity 47.1%; Pred. No. 0.078;
Matches 121; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 191 CAGGAAGGCTGTGAGTGGCTGAGATTGTCAAGCGGGTCCCAGGCGCTGCACCCAGC 250
Db 25559 CAAGGAAGCCAGGGTCCCCGGACCACCTTCTGCAACCCAGGGGTGCTGAGGTGCACCTGC 25500
QY 251 TCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACCTGACACAG 310
Db 25499 TGCCCAGGGAACCTGAGTTGAGAAGGTGGCGCCCTGCCCCATGGAAGACCAGGCCAG 25440
QY 311 GGCTAGACCCCTCACAGTGCGCCCATGTGCGCTGAGTGTGGTGTGCTGCTCAGCCGGG 370
Db 25439 AGGGTGGCCTGGGCACACTGCAGAGAGTTGCAAGAACCAAGGTTTCTGGTGGCTTGCCCCAG 25380
QY 371 ACCCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCTGGCCTGGGTT 430
Db 25379 GGCTCCTGGTCCCGGGCTTCTGGGTCCCCCTTCTCAGGTGCGCCTGGCCACCTTCTGG 25320
QY 431 CCATGCCAGCTCCAGC 447
Db 25319 CCTTCCCTGCACCAAC 25303

RESULT 15
US-10-156-761-4789/c

; Sequence 4789, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4789
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(972)
US-10-156-761-4789

Query Match 7.9%; Score 38.8; DB 15; Length 972;
Best Local Similarity 47.5%; Pred. No. 0.09;
Matches 115; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 197 AGGCTGTGAGCTGCGCTGAGATTGTCAAGCGGGGTCCCAGGCGCTGCACCGCTCACCA 256
Db 497 AGCCCGAGGGCTCGCCGAGGAGCGTGAGAGCGGTCTCCAGCCCCGGCCATCTCCGCC 438
QY 257 AGTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACCTGACACAGGGCTAG 316
Db 437 AGATACTCCTCGCACGGTTCTCGCAGCTCGGGTGCCTCGGCGGCCAGACGTTTCGGCGCG 378
QY 317 ACCCCCTCACAGTGCGCCCATGTGCCTGCAGTGTGGTGTGCTCAGCGGGACCCCC 376
Db 377 TACCACTCCGCGTTGGTCTCCGGGTCTCCGAGGCTCTCGAACGGCTCGTGCCTCGCGAACGTCAGCGAC 318
QY 377 TGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCTGGCCTGGGTTCCATGC 436
Db 317 TCCTTGAGGTGCGGCGGGGTCTCGGTGCCCTCGGAGTAGCCGTTGGCCTCGGCACCCGGA 258
QY 437 CC 438
Db 257 CC 256

Search completed: January 31, 2005, 19:08:55
Job time : 434 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:32:30 ; Search time 2949 Seconds
(without alignments)
6765.348 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgacgtgta 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 88617144

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:
17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq:
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:

44: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:
45: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:
46: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:
47: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:
48: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:
49: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:
50: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:
51: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:
52: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:
53: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:
54: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:
55: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:
56: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:
57: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:
58: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:
59: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:
60: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:
61: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:
62: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:
63: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:
64: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:
65: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:
66: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:
67: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:
68: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:
69: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:
70: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:
71: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:
72: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:
73: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:
74: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:
75: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:
76: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:
77: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:
78: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:
79: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:
80: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:
81: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:
82: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:
83: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:
84: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:
85: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:
86: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:
87: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:
88: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:
89: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:
90: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:
91: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:
92: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:
93: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:
94: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:
95: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:
96: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:
97: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:
98: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:
99: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:
100: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:
101: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:
102: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:
103: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:
104: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:
105: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:
106: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:
107: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:
108: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:
109: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:
110: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:
111: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:
112: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:
113: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:
114: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:
115: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:
116: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:

RESULT 2
PCT-US03-26864-590
; Sequence 590, Application PC/TUS0326864
; GENERAL INFORMATION:
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: METHODS OF USE FOR NOVEL POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES
; FILE REFERENCE: 08940.0013-00304
; CURRENT APPLICATION NUMBER: PCT/US03/26864
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,640
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,576
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,666
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,611
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,612
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/411,019
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,024
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,046
; PRIOR FILING DATE: 2002-09-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 731
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 590
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-26864-590

Query Match 100.0%; Score 492; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db 169 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 228

QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 229 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 289 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 348

QY 181 GGCAGGGCTGCAGGAAAGGCTGTACAGTGCAGCTGCGCTGAGATTGTCAAGCGGCGGCTCCAGGC 240
Db 349 GGCAGGGCTGCAGGAAAGGCTGTACAGTGCAGCTGCGCTGAGATTGTCAAGCGGCGGCTCCAGGC 408

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGAGCTGGGTCGAGCCTCA 300
Db 409 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGAGCTGGGTCGAGCCTCA 468

QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGCAGTGCAGCTGCGCTGAGTGTGGGTGCTG 360
Db 469 CCTGACACAGGGCTAGACCCCTTACAGTGCAGTGCAGTGCAGTGTGGGTGCTG 528

QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCT 420
Db 529 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACCCCT 588

QY 421 GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480

Db 589 GGCCTGGGTTCCATGCCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 648

QY 481 ACCCGATCGTGA 492
Db 649 ACCCGATCGTGA 660

RESULT 3
PCT-US03-27107-172
; Sequence 172, Application PC/TUS0327107
; GENERAL INFORMATION:
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: NOVEL HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES
; FILE REFERENCE: 08940.0011-00304
; CURRENT APPLICATION NUMBER: PCT/US03/27107
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,640
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,576
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,666
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,611
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,612
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/411,019
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,024
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,046
; PRIOR FILING DATE: 2002-09-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 731
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-27107-172

Query Match 100.0%; Score 492; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60
Db 169 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 228

QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 229 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 288

QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db 289 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 348

QY 181 GGCAGGGCTGCAGGAAAGGCTGTACAGTGCAGCTGCGCTGAGATTGTCAAGCGGCGGCTCCAGGC 240
Db 349 GGCAGGGCTGCAGGAAAGGCTGTACAGTGCAGCTGCGCTGAGATTGTCAAGCGGCGGCTCCAGGC 408

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGAGCTGGGTCGAGCCTCA 300
Db 409 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGAGCTGGGTCGAGCCTCA 468

QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGCAGTGCAGCTGCGCTGAGTGTGGGTGCTG 360
Db 469 CCTGACACAGGGCTAGACCCCTTACAGTGCAGTGCAGTGCAGTGTGGGTGCTG 528

QY	361	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	420
Db	529	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT	588
QY	421	GGCCTGGGTTCCATGCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC	480
Db	589	GGCCTGGGTTCCATGCCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC	648
QY	481	ACCCGATCGTGA	492
Db	649	ACCCGATCGTGA	660

RESULT 4
PCT-US03-27107-590
; Sequence 590, Application PC/TUS0327107
; GENERAL INFORMATION:
; APPLICANT: FIVEPRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: NOVEL HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES
; FILE REFERENCE: 08940.0011-00304
; CURRENT APPLICATION NUMBER: PCT/US03/27107
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/406,616
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,655
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,640
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,576
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,666
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,611
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/406,612
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/411,019
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,024
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/411,046
; PRIOR FILING DATE: 2002-09-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 731
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 590
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-27107-590

Query Match	100.0%;	Score 492;	DB 2;	Length 660;
Best Local Similarity	100.0%;	Pred. No. 2.4e-110;		
Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	60	
Db	169	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG	228	
QY	61	CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAAATTCGCAACCTGCTG	120	
Db	229	CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAAATTCGCAACCTGCTG	288	
QY	121	GGGTTGGCTCTGGGTCCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT	180	
Db	289	GGGTTGGCTCTGGGTCCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT	348	
QY	181	GGCAGGSGTGCAGGAAAGGCTGTACGCTGCCGTGAGATTGTCAAGCGGGGGTCCCAGGC	240	
Db	349	GGCAGGSGTGCAGGAAAGGCTGTACGCTGCCGTGAGATTGTCAAGCGGGGGTCCCAGGC	408	
QY	241	CTGCACCAGCTCACCAAGCTACGTTTCCCTTCAGACTAGGACAGCTGGGTCCCAGCCTCA	300	

Db	409	CTGCACCAGCTCACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA	468
QY	301	CCTGACACAGGGCTAGACCCCTCACAGTGGCGCCGCGCATGTGCCTGCAGTGTGGGTGCTG	360
Db	469	CCTGACACAGGGCTAGACCCCTCACAGTGGCGCCGCGCATGTGCCTGCAGTGTGGGTGCTG	528
QY	361	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCCCT	420
Db	529	CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCCCCT	588
QY	421	GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC	480
Db	589	GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGCTCGAGAC	648
QY	481	ACCCGATCGTGA	492
Db	649	ACCCGATCGTGA	660

RESULT 5
US-10-170-235-36232
; Sequence 36232, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES TH
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 36232
; LENGTH: 889
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-36232

[illegible]

QY 481 ACCGATCGTGA 492
| | | | |
Db 559 ACCGATCGTGA 570

RESULT 6

US-60-452-680-2683
; Sequence 2683, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2683
; LENGTH: 889
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-2683

Query Match 100.0%; Score 492; DB 111; Length 889;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 60
| | | | |
Db 79 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 138

QY 61 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
| | | | |
Db 139 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 198

QY 121 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTCTCAGGTTCT 180
| | | | |
Db 199 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTCTCAGGTTCT 258

QY 181 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGGCGGTCGCCAGGC 240
| | | | |
Db 259 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGGCGGTCGCCAGGC 318

QY 241 CTGCACCAAGCTACACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCTCA 300
| | | | |
Db 319 CTGCACCAAGCTACACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCTCA 378

QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCT 360
| | | | |
Db 379 CCTGACACAGGGCTAGACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCT 438

QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCT 420
| | | | |
Db 439 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCT 498

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
| | | | |
Db 499 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 558

QY 481 ACCGATCGTGA 492
| | | | |
Db 559 ACCGATCGTGA 570

RESULT 7

US-09-731-872-218
; Sequence 218, Application US/09731872
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 218
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..566
US-09-731-872-218

Query Match 100.0%; Score 492; DB 32; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 60
| | | | |
Db 78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCAATGCCCCAG 137

QY 61 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
| | | | |
Db 138 CTACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 197

QY 121 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTCTCAGGTTCT 180
| | | | |
Db 198 GGGTTGGCTCTGGTCCGTTGGAGGGCGGCGAGTGTCTGGCATGTAGTCTCAGGTTCT 257

QY 181 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGGCGGTCGCCAGGC 240
| | | | |
Db 258 GGCAGGGCTGCAGGAAAGCTGTCAAGTGGCTGAGATTGTCAAGCGGGCGGTCGCCAGGC 317

QY 241 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGCCAGCTGGGTCGCCAGCTCA 300
| | | | |
Db 318 CTGCACCAAGCTACCAAGCTACGTTTCTTCCAGTGGGTCGCCAGCTGGGTCGCCAGCTCA 377

QY 301 CCTGACACAGGGCTAGACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCT 360
| | | | |
Db 378 CCTGACACAGGGCTAGACCCCTTACAGTGGCGCGCATGTGCTTCCAGTGGGTCGCCAGCT 437

QY 361 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCT 420
| | | | |
Db 438 CTCAGCCGGGACCCCTTGGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCACCCT 497

QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
| | | | |
Db 498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 557

QY 481 ACCGATCGTGA 492
| | | | |
Db 558 ACCGATCGTGA 569

RESULT 8

US-09-876-997-218
; Sequence 218, Application US/09876997
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470

;
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 218
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..566
US-09-876-997-218

Query Match 100.0%; Score 492; DB 36; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 78 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137
QY 61 CTACCTCCTGATACCTTTGAGATGCGGTCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 138 CTACCTCCTGATACCTTTGAGATGCGGTCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 198 GGGTTGGCTCTGGTTCGGTTGGAGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 257
QY 181 GGCAGGGCTGCAGGAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db 258 GGCAGGGCTGCAGGAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 317
QY 241 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db 318 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 377
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGAGTGGGTGCTG 360
Db 378 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGAGTGGGTGCTG 437
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCT 420
Db 438 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCT 497
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC 480
Db 498 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC 557
QY 481 ACCCGATCGTGA 492
Db 558 ACCCGATCGTGA 569

RESULT 9
US-60-169-629-350
; Sequence 350, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 350
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

;
; NAME/KEY: CDS
; LOCATION: 78..566
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 78..233
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.7
; OTHER INFORMATION: seq LGLALGRLEGGSA/RH
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 858..863
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 878..894
US-60-169-629-350

Query Match 100.0%; Score 492; DB 81; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 60
Db 78 ATGGAGCACTACCGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCAATGCCCCAG 137
QY 61 CTACCTCCTGATACCTTTGAGATGCGGTCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 138 CTACCTCCTGATACCTTTGAGATGCGGTCGAGATGGCAGCAAAATTCGCAACCTGCTG 197
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 198 GGGTTGGCTCTGGTTCGGTTGGAGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 257
QY 181 GGCAGGGCTGCAGGAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db 258 GGCAGGGCTGCAGGAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 317
QY 241 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 300
Db 318 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGACTGAGGACAGCTGGTCCCAGCCTCA 377
QY 301 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGAGTGGGTGCTG 360
Db 378 CCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCCATGTGCTGAGTGGGTGCTG 437
QY 361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCT 420
Db 438 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGGTTACCAACCCCGAGAGCACCCCT 497
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC 480
Db 498 GGCCTGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGCTCGAGAC 557
QY 481 ACCCGATCGTGA 492
Db 558 ACCCGATCGTGA 569

RESULT 10
US-60-187-470-350
; Sequence 350, Application US/60187470
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: 78.US2.PRO
; CURRENT APPLICATION NUMBER: US/60/187,470
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 350
; LENGTH: 894
; TYPE: DNA


```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..566
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 78..233
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.7
; OTHER INFORMATION: seq LGLALGRLEGGSA/RH
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 858..863
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 878..894
US-60-187-470-350

Query Match      100.0%; Score 492; DB 83; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db      78 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 137

QY      61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACTGCTG 120
Db      138 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACTGCTG 197

QY      121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGCCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db      198 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGCCTGAGATTGTCAAGCGCGGGTCCCAGGC 257

QY      181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 317
Db      258 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 377

QY      241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db      318 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 377

QY      301 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCCTGCAGTGGGTGCTG 360
Db      378 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCCTGCAGTGGGTGCTG 437

QY      361 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGTTACCAACCCCAAGGAGCACCCT 420
Db      438 CTCAGCCGGGACCCCTGGACCCCAATGAGTGTGTTACCAACCCCAAGGAGCACCCT 497

QY      421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGTCCGAGAC 480
Db      498 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGTCCGAGAC 557

QY      481 ACCCGATCGTGA 492
Db      558 ACCCGATCGTGA 569

RESULT 11
US-09-649-163-8727
; Sequence 8727, Application US/09649163
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Galvin, Katherine A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kingsbury, Gillian A.
; APPLICANT: Weich, Nadine S.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Williamson, Mark
; APPLICANT: Richardson, Jennifer
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Fraser, Christopher C.

```

```

; APPLICANT: Villevall, Jean-Luc M.G.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; APPLICANT: Busfield, Samantha J.
; APPLICANT: Deeds, James
; APPLICANT: Lee, John
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1164-001
; CURRENT APPLICATION NUMBER: US/09/649,163
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/150,608
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 10535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8727
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-163-8727

Query Match      100.0%; Score 492; DB 28; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCAGCGCCTTCCCCAATGCCCCAG 60
Db      275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCAGCGCCTTCCCCAATGCCCCAG 334

QY      61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db      335 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394

QY      121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGCCTCGGCATGTAGTTCTCAGGTTCT 180
Db      395 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGCCTCGGCATGTAGTTCTCAGGTTCT 454

QY      181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db      455 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 514

QY      241 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db      515 CTGCACCAAGCTACCAAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 574

QY      301 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCCTGCAGTGGGTGCTG 360
Db      575 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCCTGCAGTGGGTGCTG 634

QY      361 CTCAGCCGGGACCCCTTGAGACCCCAATGAGTGTGTTACCAACCCCAAGGAGCACCCT 420
Db      635 CTCAGCCGGGACCCCTTGAGACCCCAATGAGTGTGTTACCAACCCCAAGGAGCACCCT 694

QY      421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGTCCGAGAC 480
Db      695 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAAGGGTCCGAGAC 754

QY      481 ACCCGATCGTGA 492
Db      755 ACCCGATCGTGA 766

RESULT 12
US-09-699-997-8992
; Sequence 8992, Application US/09699997
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: DiStefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

```



```
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2020-001
; CURRENT APPLICATION NUMBER: US/09/699,997
; CURRENT FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 60/162,359
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 12714
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8992
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-699-997-8992

Query Match      100.0%; Score 492; DB 30; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 334
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 335 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 395 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 454
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db 455 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 514
QY 241 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGCTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 515 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGCTGAGGACAGCTGGGTCCCAGCCTCA 574
QY 301 CCTGACACAGGGCTAGACCCCTCCAGCTGCAGTGGCGGCCCATGTGCCTGAGTGTGGTGTG 360
Db 575 CCTGACACAGGGCTAGACCCCTCCAGCTGCAGTGGCGGCCCATGTGCCTGAGTGTGGTGTG 634
QY 361 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 635 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 694
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 695 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
Db 755 ACCCGATCGTGA 766

RESULT 13
US-09-736-119-1704
; Sequence 1704, Application US/09736119
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villevall, Jean-Luc M.G.
; APPLICANT: Fraser, Christopher C.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2047-001
; CURRENT APPLICATION NUMBER: US/09/736,119
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/170,468
; PRIOR FILING DATE: 1999-12-13
; NUMBER OF SEQ ID NOS: 2118
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1704
; LENGTH: 1086
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-119-1704

Query Match      100.0%; Score 492; DB 32; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 334
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 335 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 395 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCT 454
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
Db 455 GGCAGGGCTGCAGGAAAGGCTGTCAAGTGGCTGAGATTGTCAAGCGCGGGTCCCAGGC 514
QY 241 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGCTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 515 CTGACCAAGCTACCAAGCTACGTTTCTTCCAGCTGAGGACAGCTGGGTCCCAGCCTCA 574
QY 301 CCTGACACAGGGCTAGACCCCTCCAGCTGCAGTGGCGGCCCATGTGCCTGAGTGTGGTGTG 360
Db 575 CCTGACACAGGGCTAGACCCCTCCAGCTGCAGTGGCGGCCCATGTGCCTGAGTGTGGTGTG 634
QY 361 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 420
Db 635 CTCAGCGGGACCCCTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGACACCCCT 694
QY 421 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 695 GGCCTGGTTCCATGCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
Db 755 ACCCGATCGTGA 766

RESULT 14
PCT-US02-05109-107
; Sequence 107, Application PC/TUS0205109
; GENERAL INFORMATION:
; APPLICANT: HySeq Inc.
; APPLICANT: Tang, Tom Y.
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-094-061
; CURRENT APPLICATION NUMBER: PCT/US02/05109
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/810,173
```

; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 1052
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 107
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (275)..(766)
PCT-US02-05109-107

Query Match 100.0%; Score 492; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 334
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGCAACCTGCTG 120
Db 335 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTGCGGCATGTAGTTCTCAGGTTCT 180
Db 395 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTGCGGCATGTAGTTCTCAGGTTCT 454
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGGCGGAGTGTCAAGCGGCGGTTCCCAGGC 240
Db 455 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGGCGGAGTGTCAAGCGGCGGTTCCCAGGC 514
QY 241 CTGCACAGGCTCACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 515 CTGCACAGGCTCACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 574
QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTG 360
Db 575 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTG 634
QY 361 CTCAGCCGGGACCCCTTCAGCTCCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 635 CTCAGCCGGGACCCCTTCAGCTCCAGCTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
Db 755 ACCCGATCGTGA 766

RESULT 15

US-09-810-173-107
; Sequence 107, Application US/09810173
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Chen, Rui-hong
; APPLICANT: Ma, Yunqing
; APPLICANT: Wehrman, Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Wang, Dunrui

; APPLICANT: Ghosh, Malabika J.
; APPLICANT: Dimanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/810,173
; CURRENT FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 107
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (275)..(763)
US-09-810-173-107

Query Match 100.0%; Score 492; DB 34; Length 1089;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db 275 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 334
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGCAACCTGCTG 120
Db 335 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTTCGCAACCTGCTG 394
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTGCGGCATGTAGTTCTCAGGTTCT 180
Db 395 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGAGTGTGCGGCATGTAGTTCTCAGGTTCT 454
QY 181 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGGCGGAGTGTCAAGCGGCGGTTCCCAGGC 240
Db 455 GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGGCGGAGTGTCAAGCGGCGGTTCCCAGGC 514
QY 241 CTGCACAGGCTCACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 300
Db 515 CTGCACAGGCTCACCAAGCTACGTTTCCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 574
QY 301 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTG 360
Db 575 CCTGACACAGGGCTAGACCCCTTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTG 634
QY 361 CTCAGCCGGGACCCCTTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCAGCCCCCT 420
Db 635 CTCAGCCGGGACCCCTTCGACCCCAATGAGTGTGGTTACCAACCCCGAGGAGCAGCCCCCT 694
QY 421 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 480
Db 695 GGCCTGGGTTCCATGCCAGCTCCAGCTGTGGCCCTCGTTCCCGAAGAGGGCTCGAGAC 754
QY 481 ACCCGATCGTGA 492
Db 755 ACCCGATCGTGA 766

Search completed: January 31, 2005, 18:58:39
Job time : 2955 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 12:16:25 ; Search time 85 Seconds
(without alignments)
4114.213 Million cell updates/sec

Title: US-10-057-813-13
Perfect score: 492
Sequence: 1 atggagcactaccggaagc.....ctcgagacacccgatcgtga 492

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	374.4	76.1	461	4	US-09-621-976-217	Sequence 217, App
2	236	48.0	303	4	US-09-513-999C-3603	Sequence 3603, Ap
3	74.6	15.2	1941	4	US-09-620-312D-5	Sequence 5, Appli
4	74.6	15.2	2375	4	US-09-976-594-92	Sequence 92, Appl
C 5	46	9.3	356	4	US-09-270-767-8210	Sequence 8210, Ap
C 6	46	9.3	356	4	US-09-270-767-23492	Sequence 23492, A
C 7	37.2	7.6	46819	3	US-09-453-702B-72	Sequence 72, Appl
8	35.6	7.2	22306	3	US-09-453-702B-251	Sequence 251, App
9	35.4	7.2	4257	2	US-08-690-473-1	Sequence 1, Appli
10	35.4	7.2	4257	3	US-09-259-821A-1	Sequence 1, Appli
11	35.4	7.2	4257	3	US-08-843-659-1	Sequence 1, Appli
12	35.4	7.2	4257	4	US-09-825-288A-1	Sequence 1, Appli
C 13	35.4	7.2	12001	1	US-08-458-568A-11	Sequence 11, Appl
C 14	35.2	7.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
15	34.8	7.1	1320	4	US-09-727-238-1	Sequence 1, Appli
C 16	34.8	7.1	2074	4	US-09-690-942-7	Sequence 7, Appli
17	34.6	7.0	531	4	US-09-252-991A-12044	Sequence 12044, A
C 18	34.4	7.0	1476	4	US-09-489-039A-5063	Sequence 5063, Ap
19	33	6.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli
20	33	6.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 21	32.6	6.6	378	6	5274075-3	Patent No. 5274075
22	32.6	6.6	2430	4	US-09-252-991A-12709	Sequence 12709, A
C 23	32.6	6.6	2607	4	US-09-252-991A-13249	Sequence 13249, A
24	32.4	6.6	950	4	US-09-270-767-11073	Sequence 11073, A
C 25	31.8	6.5	1950	4	US-09-252-991A-11086	Sequence 11086, A
C 26	31.8	6.5	2127	4	US-09-252-991A-8192	Sequence 8192, Ap
27	31.8	6.5	2874	4	US-09-252-991A-8112	Sequence 8112, Ap

C	28	31.6	6.4	1206	4	US-09-489-039A-3584	Sequence 3584, Ap
	29	31.6	6.4	1248	4	US-09-489-039A-3430	Sequence 3430, Ap
C	30	31.4	6.4	1204	4	US-09-620-312D-501	Sequence 501, App
	31	31.4	6.4	1232	4	US-09-270-767-11298	Sequence 11298, A
	32	31.4	6.4	2218	2	US-08-845-998-5	Sequence 5, Appli
	33	31.4	6.4	2218	3	US-09-206-537-5	Sequence 5, Appli
	34	31.4	6.4	2218	3	US-09-430-854-5	Sequence 5, Appli
C	35	31.2	6.3	3479	4	US-10-140-002-123	Sequence 123, App
C	36	31.2	6.3	8310	3	US-08-870-126-11	Sequence 11, Appl
C	37	31.2	6.3	8310	4	US-09-445-247-11	Sequence 11, Appl
C	38	31.2	6.3	14985	1	US-08-652-972A-6	Sequence 6, Appli
C	39	31.2	6.3	14985	5	PCT-US96-06231A-6	Sequence 6, Appli
C	40	31	6.3	483	4	US-09-252-991A-13386	Sequence 13386, A
	41	31	6.3	501	4	US-09-854-133-486	Sequence 486, App
	42	31	6.3	945	4	US-09-252-991A-13854	Sequence 13854, A
C	43	31	6.3	1017	4	US-09-252-991A-13594	Sequence 13594, A
C	44	31	6.3	1721	1	US-07-688-352C-13	Sequence 13, Appl
C	45	31	6.3	1721	2	US-08-474-379C-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-217
; Sequence 217, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 217
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..461
; NAME/KEY: sig_peptide
; LOCATION: 78..233
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.69999980926514
; OTHER INFORMATION: seq LGLALGRLEGGSA/RH
US-09-621-976-217

Query Match 76.1%; Score 374.4; DB 4; Length 461;
Best local Similarity 98.4%; Pred. No. 7.4e-98;
Matches 376; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAGCGCCTTCCCCAATGCCCCAG	60
Db	78	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCAGCGCCTTCCCCAATGCCCCAG	137
Qy	61	CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAATTCGCAACCTGCTG	120
Db	138	CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAATTCGCAACCTGCTG	197
Qy	121	GGTTGGCTCTGGGTGGAGGGCGGCAGTGTGCGCATGTAGTGTCTCAGGTTCT	180
Db	198	GGTTGGCTCTGGGTGGAGGGCGGCAGTGTGCGCATGTAGTGTCTCAGGTTCT	257
Qy	181	GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGTCCGAGATGGCAGCAAATTCGCAACCTG	240
Db	258	GGCAGGGCTGCAGGAAAGGCTGTGAGTGGCGTCCGAGATGGCAGCAAATTCGCAACCTG	317
Qy	241	CTGCACCAGCTCACCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA	300

Db 318 CTGCACAGCTCACCAGCTACGTTTCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCA 377
Qy 301 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCTGAGTGTGGGTGCTG 360
Db 378 CCTGACACAGGGCTAGACCCCTCAGAGTGGCGGCCATGTGCTGAGTGTGGGTGCTG 437
Qy 361 CTCAGCGGGGACCCCTGGACC 382
Db 438 CTCAGCGGGDMCCCTGGACCC 459

RESULT 2
US-09-513-999C-3603
; Sequence 3603, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3603
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..301

US-09-513-999C-3603

Query Match 48.0%; Score 236; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.6e-58;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCCTCCCAATGCCCCAG 60
Db 68 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCCTCCCAATGCCCCAG 127
Qy 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 128 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 187
Qy 121 GGGTTGGCTCTGGTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTTCTCAGGTTCT 180
Db 188 GGGTTGGCTCTGGTCGGTTGGAGGGCGGAGTCTCGGCATGTAGTTCTCAGGTTCT 247
Qy 181 GGCAGGGCTGAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCC 236
Db 248 GGCAGGGCTGAGGAAAGGCTGTGAGTGGCTGAGATTGTCAAGCGCGGGTCCC 303

RESULT 3
US-09-620-312D-5
; Sequence 5, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Changhai
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 5
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (138)..(773)
US-09-620-312D-5

Query Match 15.2%; Score 74.6; DB 4; Length 1941;
Best Local Similarity 57.5%; Pred. No. 7.8e-12;
Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 36 CCCAGCGCCTTCCCCAATGCCCCAGCTACCTCCTGATACCTTTGAGATGCGGGTCCGAGA 95
Db 239 CCGGGCTCCGGCCCGTTCGAGACCTGGCGCCGCGGTGCACATGCGGGTCAAGGA 298
Qy 96 TGGCAGCAAAATTCCGAACCTGCTGGGTTGGCTCTGGGTGGTTGGAGGCGGAGTGC 155
Db 299 AGGAGCAAGATCCGGAACCTGATGGCCTTCGCCACCGCCAGCATGGCGCAGCCAC 358
Qy 156 TCGGCATGTAGTGTCTCAGGTTCTGGCAGGGTTCGAGGAAAGGCTGTACGCTGCGCTGA 215
Db 359 GCGCGCCATCGTCTTCAGCGGCTGCGGCGGCGCCACCAACCCGTCACGTCGCGCGGA 418
Qy 216 GATGTCAAGCGCGGCTCCAGGCTGCACCAAGCTCACCAGCTACGTTTCC 268
Db 419 GATCTCAAGCGCGCCTCGGCGGCTGCACCAAGCTCAGCGGCTCGGCTACC 471

RESULT 4
US-09-976-594-92
; Sequence 92, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 92
; LENGTH: 2375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 337221.11
US-09-976-594-92

Query Match 15.2%; Score 74.6; DB 4; Length 2375;

```
Best Local Similarity 57.5%; Pred. No. 8.4e-12;
Matches 134; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 36 CCCAGCGCCTTCCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGATCGGGTCCGAGA 95
Db 271 CCCGGCTCCGGCCCTTCGCAGACCTGGCGCGGGCGGTGCACATGCGGTCAAGGA 330

QY 96 TGGCAGCAAAATTCCGAACCTGCTGGGTTGGCTCTGGGTCGGTTGGAGSGGCGCAGTGC 155
Db 331 AGGCAGCAAGATCCGGAACCTGATGGCCCTTCGCCACCGCCAGCATGGCGCAGCCACCCAC 390

QY 156 TCGGCATGTAGTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTACGTCGCTGA 215
Db 391 GCGGCCCATCGTCTTCAGCGGCTGCGGCGGCCGCCACCACCAAAACCGTCACTGCGCCGA 450

QY 216 GATTGTCAAGCGGGTCCAGGCTGCACCAAGCTCACCAGCTACGTTTCC 268
Db 451 GATCCTCAAGCGCCGCTGGCGGGCCTGCACCAAGTCAACGGCTGCGCTACC 503
```

```
RESULT 5
US-09-270-767-8210/c
; Sequence 8210, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8210
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8210
```

```
Query Match 9.3%; Score 46; DB 4; Length 356;
Best Local Similarity 51.5%; Pred. No. 0.00067;
Matches 106; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 31 GAGCTCCAGCGCCTTCCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGATCGGGTC 90
Db 354 GATCTGCCATTTCGAGGACTGCTGCAAGTCCCAAAAGTCCCAAAAGGATTTTTTGTGGATGCATGTA 295

QY 91 CGAGATGGCAGCAAAATTCCGAACCTGCTGGGTTGGCTCTGGGTCGGTTGGAGGGCGGC 150
Db 294 AAAGGCGGTACCAAGTGAGCAATGTGATTGATTGCTCAGGAGGCGCTGAACAAGGGC 235

QY 151 AGTGCTCGGCATGTAGTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTACGCTGC 210
Db 234 GAGCACAGATGCGTGTGTGGAGCGGATCCGCGGAGGAGTGGGGAAGACTATATATCGTGT 175

QY 211 GCTGAGATTGTCAAGCGCGGGTCCC 236
Db 174 GCCGAGGTCCTCAAGCGGAGCCACCC 149
```

```
RESULT 6
US-09-270-767-23492/c
; Sequence 23492, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23492
; LENGTH: 356
```

```
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-23492

Query Match 9.3%; Score 46; DB 4; Length 356;
Best Local Similarity 51.5%; Pred. No. 0.00067;
Matches 106; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 31 GAGCTCCAGCGCCTTCCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGATCGGGTC 90
Db 354 GATCTGCCATTTCGAGGACTGCTGCAAGTCCCAAAAGTCCCAAAAGGATTTTTTGTGGATGCATGTA 295

QY 91 CGAGATGGCAGCAAAATTCCGAACCTGCTGGGTTGGCTCTGGGTCGGTTGGAGGGCGGC 150
Db 294 AAAGGCGGTACCAAGTGAGCAATGTGATTGATTGCTCAGGAGGCGCTGAACAAGGGC 235

QY 151 AGTGCTCGGCATGTAGTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTGTACGCTGC 210
Db 234 GAGCACAGATGCGTGTGTGGAGCGGATCCGCGGAGGAGTGGGGAAGACTATATATCGTGT 175

QY 211 GCTGAGATTGTCAAGCGCGGGTCCC 236
Db 174 GCCGAGGTCCTCAAGCGGAGCCACCC 149
```

```
RESULT 7
US-09-453-702B-72/c
; Sequence 72, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46819
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-453-702B-72
```



```
RESULT 10
US-09-259-821A-1
; Sequence 1, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259,821A
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-259-821A-1

Query Match      7.2%; Score 35.4; DB 3; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCCCATGTGCCTGCAGTGTGGTGTGCTCAGCCGGGACCCCTGGACCCCA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGGCGGACGAGGCGCGCGTGGCGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGACACCCCTGGCCTGGGTTCCATGCCAGCTCCA 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2344 GTGAGCCTGGTCGCCGGGGCCCTGGGCCCGCGCTGCCGCGGACCCGCGCCTGCCGAGC 2403

QY 446 GCTGTGGCCCTCGTTCCCGAAGAAGGGCT 474
    ||||| ||||| ||||| ||||| |||||
Db 2404 TCCGCGCGCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 11
US-08-843-659-1
; Sequence 1, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Roasrio
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARSB:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-843-659-1

Query Match      7.2%; Score 35.4; DB 3; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCCCATGTGCCTGCAGTGTGGTGTGCTCAGCCGGGACCCCTGGACCCCA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGGCGGACGAGGCGCGCGTGGCGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGACACCCCTGGCCTGGGTTCCATGCCAGCTCCA 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2344 GTGAGCCTGGTCGCCGGGGCCCTGGGCCCGCGCTGCCGCGGACCCGCGCCTGCCGAGC 2403

QY 446 GCTGTGGCCCTCGTTCCCGAAGAAGGGCT 474
    ||||| ||||| ||||| ||||| |||||
Db 2404 TCCGCGCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 12
US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Patent No. 6723511
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match      7.2%; Score 35.4; DB 4; Length 4257;
Best Local Similarity 52.3%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 326 CAGTGGCGCCCATGTGCCTGCAGTGTGGTGTGCTCAGCCGGGACCCCTGGACCCCA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2284 CTGCGGGGACCTGCGGTGGCGGGCGGACGAGGCGCGCGTGGCGCGCC 2343

QY 386 ATGAGTGTGTTACCAACCCCGAGGACACCCCTGGCCTGGGTTCCATGCCAGCTCCA 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2344 GTGAGCCTGGTCGCCGGGGCCCTGGGCCCGCGCTGCCGCGGACCCGCGCCTGCCGAGC 2403

QY 446 GCTGTGGCCCTCGTTCCCGAAGAAGGGCT 474
    ||||| ||||| ||||| ||||| |||||
Db 2404 TCCGCGCGCGCGCGCGCGCGGACCTGCT 2432

RESULT 13
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
```


Db 1089 CAACGTCTACGAGCGCC 1106

Search completed: January 31, 2005, 18:09:29
Job time : 98 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:01 ; Search time 32.6 Seconds
(without alignments)
481.084 Million cell updates/sec

Title: US-10-057-813-14

Perfect score: 861

Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....SMPSSCGPRRRRRARDTRS 163

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	219.5	25.5	538 2 H86335	T20H2.2 protein -
2	105.5	12.3	369 2 F96788	protein T4O12.22 [
3	86.5	10.0	350 2 C26984	probable regulator
4	85.5	9.9	906 2 T00039	hypothetical prote
5	84.5	9.8	368 2 T35150	probable glycosyl
6	84	9.8	397 2 T30168	hypothetical prote
7	84	9.8	634 2 T00388	hypothetical prote
8	82.5	9.6	556 2 D70940	probable PPE prote
9	81.5	9.5	1275 2 AD0332	probable membrane
10	81	9.4	367 2 C39590	TPA-induced protei
11	80.5	9.3	257 2 T10586	small nuclear ribo
12	80.5	9.3	495 1 P2WLB2	L2 protein - bovin
13	80	9.3	589 2 C70767	probable pknJ - My
14	79.5	9.2	1052 2 T00067	hypothetical prote
15	79	9.2	469 1 P2WLB	L2 protein - bovin
16	79	9.2	770 2 G70718	probable cation tr
17	78.5	9.1	359 2 T36443	probable penicilli
18	78.5	9.1	579 2 C75599	TerF-related prote
19	78.5	9.1	2605 2 T18552	saframycin Mx1 syn
20	78	9.1	817 2 S51342	verprolin - yeast
21	78	9.1	827 2 H83217	probable transcrip
22	77.5	9.0	6420 2 T30283	polyketide synthas
23	76.5	8.9	467 2 T34874	hypothetical prote
24	76.5	8.9	474 2 T19543	hypothetical prote
25	76.5	8.9	1460 1 EDBEIF	immediate-early pr
26	76.5	8.9	2142 2 B35098	MHC class III hist
27	76	8.8	202 2 AF2088	hypothetical prote
28	76	8.8	434 1 WMBEF3	UL43 protein - hum
29	76	8.8	537 2 I53719	NF-kappa-B transcr

30	76	8.8	2774	2 A43359	microtubule-associ
31	76	8.8	3343	2 T42207	breast cancer susc
32	75.5	8.8	231	2 D72532	hypothetical prote
33	75.5	8.8	405	2 T44488	araC-family transc
34	75.5	8.8	433	2 F75566	probable benzoate
35	75.5	8.8	466	2 AH3100	amidohydrolase [im
36	75.5	8.8	466	2 B98186	probable hydrolase
37	75	8.7	271	2 S48426	RHR2 protein - yea
38	75	8.7	278	2 T31319	triose-phosphate i
39	75	8.7	307	2 B72677	hypothetical prote
40	75	8.7	560	2 D30930	flagellar basal bo
41	75	8.7	565	2 T08794	hypothetical prote
42	75	8.7	1446	1 A45344	immediate-early pr
43	75	8.7	1493	2 A38218	GAP-associated pro
44	75	8.7	7463	2 T36248	CDA peptide synthe
45	74.5	8.7	336	2 T17408	rRNA (adenine-N6-)

ALIGNMENTS

RESULT 1

H86335

T20H2.2 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: H86335

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H86335

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-538 <STO>

A;Cross-references: UNIPROT:Q9LNU5; GB:AE005172; NID:g8778978; PIDN:AAF79893.1; GSPDB:G C;Genetics:

A;Map position: 1

Query Match 25.5%; Score 219.5; DB 2; Length 538;

Best local Similarity 30.4%; Pred. No. 4.6e-12;

Matches 55; Conservative 31; Mismatches 64; Indels 31; Gaps 6;

QY 1 MEHYRKAGSVELP-APSPMPQLPPDITLEMVRDGSKIRNLLGLALGRLEGSSARHVVFSG 59

Db 224 MDKYQR---VEKPKADTPIAE-----NEIRITSMGRARNYITYAMALLQENKSNVEIFKA 275

QY 60 SGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGTGLDPLTVRRHVPAVMV 119

Db 276 MGRAINKSVTIVELIKRIPGLHQITSIGSTDITDTWEPT--EEGLQTIETTRHVSMTIT 333

QY 120 LLSRDPLDPNECGYQPP-----GAPPLGLSMPSSSCGPRRRRRARDT 161

Db 334 TLSKEQLNTSSVGVCQPIPIEMVKPLAEIDYEGQDGSPRGRGRRGR--GGRGRGRGG 391

QY 162 R 162

Db 392 R 392

RESULT 2

F96788

protein T4O12.22 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F96788

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96788
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-369 <STO>
A;Cross-references: UNIPROT:Q9LQR7; GB:AE005173; NID:g8778814; PIDN:AAF79819.1; GSPDB:GN C;Genetics:
A;Gene: T4O12.22
A;Map position: 1

Query Match	12.3%	Score 105.5;	DB 2;	Length 369;
Best Local Similarity	22.5%	Pred. NO. 0.055;		
Matches 34;	Conservative 17;	Mismatches 47;	Indels 53;	Gaps 4;

Qy	1	MEHYRKAGSVELPAPSPMPQLP	PPDTTLEMRVRD	SGKIRNLLGALGRLE	-----	48
		: ::		::	::	:
Db	1	MDKYQR	-----	VVKPKADTPIDANEIRITSQGRARNYITAMTLLQVFNES	EMCQFTI	53
Qy	49	-----	GG SARHVVFGSG	RAAGKAVSCAEIVK	-----	76
			:			::
Db	54	HCLVFVACWEI	FRDKGSTE	VVFVKAMGRAINKTTIVELIKVF	FWRFALSNRSRV	QMGR
Qy	77	RVPGLHLQTKL	RFLQTEDSWVP	ASEDDTGLDP		107
		:		:	::	
Db	114	RIPDLHONT	SIGSTDITDTWEPT	--EEGLLP		142

RESULT 3
C26984
probable regulatory protein 5 - Streptomyces griseus
C/Species: Streptomyces griseus
C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C/Accession: C26984
R/Distler, J.; Ebert, A.; Mansouri, K.; Pissowotzki, K.; Stockmann, M.; Piepersberg, W.
Nucleic Acids Res. 15, 8041-8056, 1987
A/Title: Gene cluster for streptomycin biosynthesis in Streptomyces griseus: nucleotide
A/Reference number: A93676; MUID:88040426; PMID:3118332
A/Accession: C26984
A/Molecule type: DNA
A/Residues: 1-350 <DIS>
A/Cross-references: UNIPROT:P08076; GB:Y00459; GB:S55493; NID:g1621271; PIDN:CAA68515.1
C/Genetics:
A/Gene: strR
C/Keywords: transcription regulation

Query Match 10.0%; Score 86.5; DB 2; Length 350;
Best Local Similarity 29.1%; Pred. NO. 2.7;
Matches 46; Conservative 11; Mismatches 52; Indels 49; Gaps 9;

Qy	1	MEHYR--KAGSVELPAPSPMPQLPPDTLEMRVRD-----SKIRNLLGLALGRLEGSA	52
Db	54	VEHIRTLAASGAELPAIVMP-----TTKRVIDGMHRLRATKMRGATEIAVRYFEGGEE	107
Qy	53	RHVFGSGRAAGKAVSCAEIVKRRVP-GL-----HQLTKLRFLOQTEDSVFPASPD--T	103
Db	108	EAFIFA-----VKSNTVTHGLPLSLDRKAAATRVLETHPSWSRAIGLAT	152
Qy	104	GLDPLTV-----RRHVPAAVWVLLSRD-----PLDPNE	130
Db	153	GLSAKTIVGTLRSCSTAGVPQSNVRIGRDGRARPLDPE	190

RESULT 4

T00039
hypothetical protein KIAA0290 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C;Accession: T00039
R;Ohara, O.; Nagase, T.; Ishikawa, K.; Nakajima, D.; Ohira, M.; Seki, N.; Nomura, N.
submitted to the EMBL Data Library, August 1997
A;Description: Prediction of the coding sequences of unidentified human genes.
A;Reference number: Z14073
A;Accession: T00039
A;Status: preliminary; translated from GB/EMBL/DBSJ
A;Molecule type: mRNA
A;Residues: 1-906 <OHA>
A;Cross-references: UNIPROT:O14526; EMBL:AB006628; NID:d1170681; PIDN:BAA22959.1; PID:d1170681
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0290

Query Match 9.9%; Score 85.5; DB 2; Length 906;
Best Local Similarity 26.5%; Pred. NO.9.2;
Matches 49; Conservative 17; Mismatches 60; Indels 59; Gaps 11;

QY	13	PAPSPMPQLPPD--TLEMVRVDGSKIRNLLGLALGRLEGS-ARHVVFGSG-----RAA	64
Dd	380	PAPARAPACSEAAAAQLRATAGS-----LILPPGPGTMKRHSRRDAAGKQPQRSA	432
QY	65	GKAVSCAEIVKRRVPGLHQLTK-----LRFLQTEDSVWPASP--	101
Dd	433	PRTSSCAE---RLQSEEQVSKNFLFGPPLESADFHDFTGSSSLGFTSSPFPSSSSPEN	488
QY	102	--DTGLDPLTVRRHV-----PAWVLLSRDPLDPNECYQPPGA----PPGLGMP---S	147
Dd	489	VEDSGLDSPS---HAAPGSPDSWVPRPGTPQSPPSCRAPPPEARGIRAPLPDSPQPLA	545
QY	148	SSCGP	152
Dd	546	SSPGP	550

RESULT 5

probable glycosyl transferase - Streptomyces coelicolor (fragment)
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T35150
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1998
 A:Reference number: Z21570
 A:Accession: T35150
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-368 <SEE>
 A:Cross-references: UNIPROT:Q8CJM1; EMBL:AL031107; PIDN:CAA19930.1; GSPDB:G
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC5A7.01

Query Match 9.8%; Score 84.5; DB 2; Length 368;
Best Local Similarity 28.0%; Pred. No. 4.2;
Matches 46; Conservative 11; Mismatches 54; Indels 53; Gaps 11;

Qy	12	LPAPSPMPQLPPD	TLEMRV	RDGSKIRN	LLGLALGR	LEGGSARH	VFSGSRAAGKAV	SCA	71
Db	47	LPVPGDWPR--	PD-----	DAARAELAR	VLRLPDGAV--	VLLDG-----	LVACG		86
Qy	72	--EIVKRRVPGL	HQLTKL	RFLQTEDS	WVPASPD	TGLDPL-----	TVRHVP	PAV--	117
Db	87	VPEVV--VPEA	ERLMAVL	VH-----	LPLGDETGLD	PAVAAGLDARERT	VLRAVPA	VA	138
Qy	118	---WV---LLSR	DPDNP	ECGYQPPGA-----	PPGLGSM	PSSSC		150	
Db	139	TSDWAVRR	LVSHHGLD	PGRVHVAAP	GADIAP	LAPAGT	DGVSK	LVC	182

RESULT 6
T30168
hypothetical protein F26B1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30168
R;Latreille, P.; Wamsley, P.; Kramer, J.
submitted to the EMBL Data Library, November 1996
A;Description: The sequence of C. elegans cosmid F26B1.
A;Reference number: Z20748
A;Accession: T30168
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-397 <LAT>
A;Cross-references: UNIPROT:P91277; EMBL:U80444; PIDN:AAB37791.1; GSPDB:GN00019; CESP:F2
A;Experimental source: strain Bristol N2; clone F26B1
C;Genetics:
A;Gene: CESP:F26B1.2
A;Map position: 1
A;Introns: 16/3; 40/2; 78/3; 93/2; 117/3; 320/3; 382/2

Query Match 9.8%; Score 84; DB 2; Length 397;
Best Local Similarity 26.5%; Pred. No. 5.1;
Matches 35; Conservative 19; Mismatches 62; Indels 16; Gaps 4;

QY 31 RDGSKIRNLLGLALGRLEGG SARHVVFGSGRAAGKAVSCAEIVKRRVPG LHQ--LTKLR 88
Db 142 RNGSKIKEL-----REKCSARLKIFTGCAPGSTDRVLITSGEQKNVLGIEEVMKELK 194
QY 89 FLQTEDSWVPASPD TGLDPLTVRRH-----VPAVWVLLSRDPLDPNECGYQPPGAP PGL 142
Db 195 EIPIKGSATPYLP AFNYDPSNISDYGGFFGNMPAGPPNNRGPA-PQ RGGGQPPGGPRSY 253
QY 143 G SMPSSSCGPRS 154
Db 254 GGAITQGGGQRS 265

RESULT 7
T00388
hypothetical protein KIAA0616 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00388
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-634 <ISH>
A;Cross-references: UNIPROT:O75114; EMBL:AB014516; NID:g3327045; PIDN:BAA31591.1; PID:g3
A;Experimental source: brain
C;Genetics:
A;Gene: KIAA0616

Query Match 9.8%; Score 84; DB 2; Length 634;
Best Local Similarity 22.0%; Pred. No. 8.5;
Matches 37; Conservative 23; Mismatches 76; Indels 32; Gaps 5;

QY 9 SVELPAPSPMPQLP PDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGSGRAAGKAV 68
Db 251 NIHFPSPLPTPLDP EEPFPALSSSSSTGNL-----AANLTHLGIGGACQGMSTPG 301
QY 69 SC AEIVKRRVPG LHLTKLRFLOTEDSWVPASPD T--GLDPLTVRRHVP AVWVLLS--- 122
Db 302 SSPQHRPAGVSP LSLSTEARRQQA SPTLSPLSPITQAVAMDALSLEQQLPYAFFTQAGSQ 361
QY 123 -----RDPLDPNECGYQPP-----GAPPLGSMPPSSC--GPR 153
Db 362 QPPPPQPPPPPPPPASQQPPPPPPPPQAPVRLPPGGLLPSPASLTRGPQ 409

RESULT 8
D70940
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: D70940
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70940
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-556 <COL>
A;Cross-references: UNIPROT:O53678; GB:AL021929; GB:AL123456; NID:g3242291; PIDN:CAA173
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: PPE

Query Match 9.6%; Score 82.5; DB 2; Length 556;
Best Local Similarity 26.0%; Pred. No. 10;
Matches 38; Conservative 20; Mismatches 47; Indels 41; Gaps 9;

QY 9 SVELPAPSPMPQLP PDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGSGRAAGKAV 68
Db 396 AVTAPTPPGPPPPV TAPPV-TGAGIQS-FGYLVGDL-----NSAAQARKAV 442
QY 69 SC AEIVKRRVPG LHLTKLRFLOTEDSWVP--ASPD TGLDPLTVRRHVP AV-----WV 119
Db 443 GTG--VRKKT P-----EPDSAEPAAAAAPEEQVP--QRRRRPKIKQLGRGYE 487
QY 120 LLSRDPLDPNECGYQPPGAPPLGSM 145
Db 488 YLDLDP----ETGHDPTGSPQGAGTL 509

RESULT 9
AD0332
probable membrane protein YPO2724 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD0332
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0332
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1275 <KUR>
A;Cross-references: UNIPROT:Q8ZD68; GB:AL590842; PIDN:CAC92963.1; PID:g15980702; GSPDB:
C;Genetics:
A;Gene: YPO2724

Query Match 9.5%; Score 81.5; DB 2; Length 1275;
Best Local Similarity 36.1%; Pred. No. 30;
Matches 22; Conservative 11; Mismatches 15; Indels 13; Gaps 3;

QY 72 EIVKRRVPG L-----HQLTK-----LRFLOTEDSWVPA-SPDTGLDPLTVRRHVP AVW 118
Db 779 ELLHEGIPGLTYAGYHEVVKKLTPLMLLQBEDRWMAQSLSSSLDPLTLRQDVLVMY 838
QY 119 V 119
Db 839 L 839

RESULT 10
C39590
TPA-induced protein 11D - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 09-Jul-2004
C;Accession: C39590
R;Varnum, B.C.; Ma, Q.; Chi, T.; Fletcher, B.; Herschman, H.R.
Mol. Cell. Biol. 11, 1754-1758, 1991
A;Title: The Tis11 primary response gene is a member of a gene family that encodes prote
A;Reference number: A39590; MUID:91141531; PMID:1996120
A;Accession: C39590
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-367 <VAR>
A;Cross-references: UNIPROT:P23949; GB:M58564; NID:g202066; PIDN:AAA72946.1; PID:g202067
C;Keywords: DNA binding; zinc finger

Query Match 9.4%; Score 81; DB 2; Length 367;
Best Local Similarity 37.5%; Pred. No. 8.7;
Matches 21; Conservative 6; Mismatches 25; Indels 4; Gaps 1;
QY 97 VPASPDGTGLDPLTVRRHVPVAVVLLSRDPLDNECGYQPPGAPGLGSMPSSSCGP 152
Db 11 VAAAPSSFTPGFLRRHSASNLHALAHPVSPGSCSPKFPGAPNGG----SSCGP 62

RESULT 11
T10586
small nuclear ribonucleoprotein-associated protein homolog F9F13.90 - Arabidopsis thalia
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10586
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16991
A;Accession: T10586
A;Molecule type: DNA
A;Residues: 1-257 <BEV>
A;Cross-references: UNIPROT:Q9SUN5; EMBL:AL080253; GSPDB:GN00062; ATSP:F9F13.90
A;Experimental source: cultivar Columbia; BAC clone F9F13
C;Genetics:
A;Gene: ATSP:F9F13.90
A;Map position: 4
C;Superfamily: proline-rich protein

Query Match 9.3%; Score 80.5; DB 2; Length 257;
Best Local Similarity 23.0%; Pred. No. 6.6;
Matches 40; Conservative 16; Mismatches 55; Indels 63; Gaps 7;
QY 20 QLPPDTLEMRVDRGSKIRNLLGLAL-----CRLEGGSARHVVFGS 60
Db 50 KLPP-AKGKKINEERDRRTLGLVLRGEEVISMVTEGPPPPPEESRAKAGSAAAVAGPGI 108
QY 61 GRAAGKAVSCAEIVKRRVPGLH-----QLTKLRFLOTEDSWVP 98
Db 109 GRAAGRGVETGPLVQAQ-PGLSGPVRGVGGPAPGMMQPIISRPPQLSAPPIIRPPGQMLP 167
QY 99 ASPDGTGLDPLTVRRHVPVAVVLLSRDPLDNECGYQPP-----GAPPGLSMP 146
Db 168 PPPFGGQGP-----PMGRGP--PPPYGMRPPPPQFSGPPPPYQGRP 207

RESULT 12
P2WLB2
L2 protein - bovine papillomavirus type 2
C;Species: bovine papillomavirus type 2
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C;Accession: B31169
R;Groff, D.E.; Mitra, R.; Lancaster, W.D.
submitted to GenBank, May 1988
A;Reference number: A94519
A;Accession: B31169

A;Molecule type: DNA
A;Residues: 1-495 <GRO>
A;Cross-references: UNIPROT:P06457; GB:X01768; GB:M24326; NID:g60859
C;Superfamily: papillomavirus L2 protein
C;Keywords: glycoprotein; late protein
F;28,459,478/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 9.3%; Score 80.5; DB 1; Length 495;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 38; Conservative 21; Mismatches 62; Indels 39; Gaps 6;
QY 4 YRKAGSVELPAPSPMPQLPPDILEMRVDRGSKIRNLLGLAL-----GRL-EGG 50
Db 44 YRTCKQAGTCPPDVIPKVEGTI-----ADKILKGLALYLGGLGIGTWSTGRVAAGG 97
QY 51 SARHVVFGSG-----RAAGKAVSCAEIVKRRVPG-----LHQLTKLRFLOTEDSWVPAS 100
Db 98 SPRIVPLRTSGSTTSLASVGSRAGAATGTRSSITGIPLDTLETIGALRPGAYEDTVLPEA 157
QY 101 P-----DTGLDPLTVRRHVPVAVVLLSRDPLDPNE 130
Db 158 PAIVTPDAVPADTGIDGLSIGTSDSTETLITLLEPEGPED 197

RESULT 13
C70767
probable pknJ - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70767
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70767
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-589 <COL>
A;Cross-references: UNIPROT:Q10697; GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98200
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: pknJ
C;Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific protein ki
F;13-276/Domain: protein kinase homology <KIN>

Query Match 9.3%; Score 80; DB 2; Length 589;
Best Local Similarity 25.6%; Pred. No. 18;
Matches 44; Conservative 10; Mismatches 68; Indels 50; Gaps 5;
QY 23 PDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGSGRA--AGKAVSCAEIVKRRVPG 80
Db 185 PEVLAGQGFDGRADLYSLGICALFRLLTGEAPFAAGAGAAVAVVAGHLHQPPPTVSDRVPG 244
QY 81 L-----HQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRH 113
Db 245 LSAAMDVAVIATAMAKDPMRFRFTSAGEFAHAAAALYGATDGVVPPSP----- 292
QY 114 VPAVWVLLSRDPLDNECGYQPPGA-----PPGLGSMPSSSCGPRRRRAR 159
Db 293 APHV---ISQGAVPGSPWQHVPVGSVTALATPPGHGWFPGLPLPRRRRYR 341

RESULT 14
T00067
hypothetical protein KIAA0453 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C;Accession: T00067
R;Seki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.; DNA Res. 4, 345-349, 1997

A;Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A;Reference number: Z14085; MUID:98116662; PMID:9455484
A;Accession: T00067
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1052 <SEK>
A;Cross-references: UNIPROT:Q9UIM0; EMBL:AB007922; NID:d1225331; PIDN:BAA32298.1
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0453

Query Match	9.2%;	Score 79.5;	DB 2;	Length 1052;
Best Local Similarity	25.2%;	Pred. No. 37;		
Matches	40;	Conservative 20;	Mismatches 36;	Indels 63; Gaps 8;

QY	19	PQLPP-DTLEMRVRDGSKI-----RNLGLA-----LGRLEGG SARHVVFS	58
Db	286	PTLPPFITLTVKGAGSEINCMNDFQDNRQLYENFIYAATYTFSGLQEGTGRPV---	342

QY	59	GSGRAAGKAVSCAEIV-----KRRVPG-----LHQLTKLRFLOTEDSWVPASPD	102
Db	343	-----ASNKAITCAELVLDVSPKTQRVILKKEPGKRSQLWRMTGTGMLAHEGSSVPHNP	398

QY	103	-----TGLDPLTVRRHVPVAVWVLLSRDP	125
Db	399	KPSAARSTEGSAILDIAGLAATVDRYEP-----LMLRKP	433

RESULT 15
P2WLB
L2 protein - bovine papillomavirus type 1
C;Species: bovine papillomavirus type 1
C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 09-Jul-2004
C;Accession: A03653
R;Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H.
Nature 299, 529-534, 1982
A;Title: The primary structure and genetic organization of the bovine papillomavirus type 1
A;Reference number: A93289; MUID:83012974; PMID:6289124
A;Accession: A03653
A;Molecule type: DNA
A;Residues: 1-469 <CHE>
A;Cross-references: UNIPROT:P03109; GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:g609
C;Superfamily: papillomavirus L2 protein
C;Keywords: late protein

Query Match	9.2%;	Score 79;	DB 1;	Length 469;
Best Local Similarity	24.4%;	Pred. No. 17;		
Matches	43;	Conservative 19;	Mismatches 64;	Indels 50; Gaps 8;

QY	4	YRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLAL-----GRL-EGG	50
Db	16	YRTCKQAGTCPPDVIPKVEGDTI-----ADKILKFGGLAIYLGGLIGIGTWSTGRVAAGG	69

QY	51	SARHVVFSGSGRAAGKAVSCAEIVKRRV-----PGLHQLTKLRFLOT-----EDSW	96
Db	70	SPRYTPL-----RTAGSTSSLASIGSRVAVTAGTRPSIGAGIPLDTLETLGALRPGVYEDTV	125

QY	97	VPASP-----DTGLDPLTVRRHVPVAVWVLLSRDPLDPNECG---YQPPGAP	139
Db	126	LPEAPAIVTPDAVPADSGLDALSIGTSDSTETLITLLEPEGPEDIAVLEQLPLDRP	181

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:40 ; Search time 174.603 Seconds
(without alignments)
537.138 Million cell updates/sec

Title: US-10-057-813-14
Perfect score: 861
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....SMPSSSCGPRRRRRARDTRS 163

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	861	100.0	163	2	Q8N5L8	Q8n5l8 homo sapien
2	815	94.7	163	2	Q99JH1	Q99jh1 mus musculu
3	814	94.5	163	2	Q9CYX4	Q9cyx4 mus musculu
4	421	48.9	224	2	Q6DGS1	Q6dgs1 brachydanio
5	350	40.7	199	2	Q91WE3	Q91we3 mus musculu
6	343	39.8	199	2	Q9BUL9	Q9bul9 homo sapien
7	339	39.4	199	2	Q9NX88	Q9nx88 homo sapien
8	263	30.5	215	2	Q7PTH5	Q7pth5 anopheles g
9	235.5	27.4	206	2	Q9V9B6	Q9v9b6 drosophila
10	230.5	26.8	206	2	Q6NNC7	Q6nnc7 drosophila
11	230.5	26.8	206	2	AAR96155	Aar96155 drosophil
12	219.5	25.5	315	2	Q944A2	Q944a2 arabidopsis
13	219.5	25.5	538	2	Q9LNU5	Q9lnu5 arabidopsis
14	206	23.9	350	2	Q93VA8	Q93va8 arabidopsis
15	201.5	23.4	732	2	Q7XJ13	Q7xj13 oryza sativ
16	200	23.2	262	2	Q7XT90	Q7xt90 oryza sativ
17	197.5	22.9	272	2	Q8H7C6	Q8h7c6 arabidopsis
18	192.5	22.4	485	2	Q9M8Z6	Q9m8z6 arabidopsis
19	167	19.4	241	2	Q7R6W6	Q7r6w6 plasmodium
20	165	19.2	248	2	Q8IAX8	Q8iax8 plasmodium
21	117	13.6	211	2	Q8IDN4	Q8idn4 plasmodium
22	108	12.5	202	2	Q7RPX4	Q7rpx4 plasmodium
23	105.5	12.3	369	2	Q9LQR7	Q9lqr7 arabidopsis
24	104	12.1	206	2	Q7YWL9	Q7ywl9 caenorhabdi
25	100.5	11.7	1080	2	Q7RVU1	Q7rvu1 neurospora
26	100	11.6	1091	2	Q9C246	Q9c246 neurospora
27	91.5	10.6	275	2	Q9AWP6	Q9awp6 oryza sativ
28	91.5	10.6	341	2	Q8ISG7	Q8isg7 stylonychia
29	91	10.6	691	2	Q7ZUE6	Q7zue6 brachydanio
30	90.5	10.5	443	2	Q9FBM9	Q9fbm9 streptomyce
31	89.5	10.4	218	2	Q9FVZ8	Q9fvz8 oryza sativ

32	88.5	10.3	215	1	CNO_MOUSE	Q8ved2 mus musculu
33	88	10.2	479	2	Q6ZAY3	Q6zay3 oryza sativ
34	88	10.2	479	2	BAC99602	Bac99602 oryza sat
35	88	10.2	525	2	Q94JH5	Q94jh5 oryza sativ
36	87.5	10.2	510	2	Q8BL39	Q8bl39 mus musculu
37	87.5	10.2	625	2	Q93H39	Q93h39 streptomyce
38	87.5	10.2	782	1	Q93H39	Q93h39 streptomyce
39	87	10.1	416	2	Q9KZE7	Q9kze7 streptomyce
40	87	10.1	711	2	Q6KAL9	Q6kal9 mus musculu
41	87	10.1	711	2	BAD21438	Bad21438 mus muscu
42	87	10.1	1098	2	Q811E7	Q811e7 mus musculu
43	86.5	10.0	350	1	STRR_STRGR	P08076 streptomyce
44	86.5	10.0	806	2	Q8VD37	Q8vd37 mus musculu
45	86.5	10.0	885	2	Q8KY52	Q8ky52 azospirillu

ALIGNMENTS

RESULT 1

Q8N5L8
ID Q8N5L8 PRELIMINARY; PRT; 163 AA.
AC Q8N5L8;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Chromosome 9 open reading frame 23 protein.
GN Name=C9orf23;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032136; AAH32136.1; -.
DR Genew; HGNC:19909; C9orf23.
DR InterPro; IPR011574; Alba DUF78.
DR ProDom; PD010497; Alba DUF78; 1.
SQ SEQUENCE 163 AA; 17631 MW; 442C8727191A0BCE CRC64;

Query Match 100.0%; Score 861; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.5e-66;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRDGSKIRNLLGLALRLEGCSARHVFFSGS 60
|||||

```
Db 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFSGS 60
Qy 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Qy 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRSRRRARDTRS 163
Db 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRSRRRARDTRS 163

RESULT 2
Q99JH1 ID Q99JH1 PRELIMINARY; PRT; 163 AA.
AC Q99JH1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein (RIKEN cDNA 2810432D09).
GN Name=2810432D09Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Ievolella C., Zara I., Lanfranchi G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3;
RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by
RC biopsy.;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3;
RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by
RC biopsy.;
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ293897; CAC34588.1; -
DR EMBL; BC047068; AAH47068.1; -
DR MGD; MGI:1917211; 2810432D09Rik.
DR InterPro; IPR011574; Alba DUF78.
DR ProDom; PD010497; Alba DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 163 AA; 17675 MW; 87675201AF87F5B6 CRC64;

Query Match 94.7%; Score 815; DB 2; Length 163;
Best Local Similarity 94.5%; Pred. No. 1.4e-62;
Matches 154; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFSGS 60
Db 1 MEQYRRAGSVELPASSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSTRHVVFSGS 60
Qy 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSWVPASPDGLDPLTVRRHVPVAVWL 120
Qy 121 LSRDPLDPNECGYQPPGAPPGGLGSMPPSSSCGPRSRRRARDTRS 163
Db 121 LSRDPLDPSECGYQPPGAPPGGLGSIPIPSGPRRRRRARDTRS 163

RESULT 3
Q9CYX4 ID Q9CYX4 PRELIMINARY; PRT; 163 AA.
AC Q9CYX4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2810432D09 product:hypothetical protein, full
DE insert sequence.
GN Name=2810432D09Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RC MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RC MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RC The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RC MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RC MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitzunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
```


RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Muramatsu M., Hayashizaki Y.;
RA Konno H., Okazaki Y., Muramatsu N., Sugahara Y., Shibata K., Itoh M.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016085; AAH16085.1; -.
DR EMBL; AK082822; BAC38636.1; -.
DR MGD; MGI:2143151; AI851155.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21037 MW; B9DE105BE3293D64 CRC64;

Query Match 40.7%; Score 350; DB 2; Length 199;
Best Local Similarity 43.6%; Pred. No. 2.6e-22;
Matches 75; Conservative 25; Mismatches 54; Indels 18; Gaps 3;
QY 1 MEHYRKAGSVELPA-----PSPMPQLPDDTLEMRVDRGSKIRNLLGLALGRLEGG 50
||:||||| ||| ||| ||| : ||:||||||| :
Db 1 MENFRKVRSEEPAGDGEKSPNSGPFADLAPGAVHMRVKEGSKIRNLLAFATASMAQP 60
QY 51 SARHVVFSGSGRAAGKAVSCAEIVKRRVPLHQLTKLRLFLQTEDSW--VPASPDTLGLDP- 107
: ||:||||| ||| |||:||||:||||:||||:||||: |||: ||| : ||| : |||
Db 61 ATRAIVFSGCGRATTKVTCAEILKRRLAGLHQVTRLRYRSRVREVWQSLPPGTPPGQTPTS 120
QY 108 -----LTVRRRHVPAVWVLLSRDLPDNECGYQPPGAPPGGLGSMPSSSCGPRS 154
||:||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 121 DPAASLSVLKNVPSLAILLSKDALDPRQLGYQPPNLSPGPSSPPTVSTSKRS 172
RESULT 6
Q9BUL9
ID Q9BUL9 PRELIMINARY; PRT; 199 AA.
AC Q9BUL9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE RPP25 protein (RNase P protein subunit p25).
GN Name=RPP25;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC MEDLINE=2198191; PubMed=12003489;
RA Guerrier-Takada C., Eder P.S., Gopalan V., Altman S.;
RT "Purification and characterization of Rpp25, an RNA-binding protein
RT subunit of human ribonuclease P.";
RL RNA 8:290-295(2002).
DR EMBL; BC002497; AAH02497.1; -.
DR EMBL; BC007270; AAH07270.1; -.

```
DR EMBL; AY034074; AAK54443.1; -.
DR Interact; Q9BUL9; -.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20632 MW; 9A4494F0297B2A81 CRC64;

Query Match
Best Local Similarity 39.8%; Score 343; DB 2; Length 199;
Matches 73; Conservative 26; Mismatches 55; Indels 18; Gaps 3;

QY 1 MEHYRKAGSVLPA-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
   ||:|||||
Db 1 MENFRKVRSEEPAGCGAEGGPGGFADLAPGAVHMRVKEGSKIRNLMAFATASMAQP 60
   ||:|||||

QY 51 SARHVFSGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSW--VPASPDGLDP- 107
   :||:|||||
Db 61 ATRAIVFGCGRATTKTVTCAEILKRRLAGLHQVTRLRYRSRVREVMQSLPPGPTQGQTPG 120

QY 108 -----LTVRRHVPVAVVLLSRDPLDPNECGYQPPGAPGLGSMPSSSCGPRS 154
   ||:|||||
Db 121 EPAASLSVLKNVPGLAILLSKDALDPRQPGYQPPNPHPGPSPPPAAPASKRS 172

RESULT 7
Q9NX88
ID Q9NX88 PRELIMINARY; PRT; 199 AA.
AC Q9NX88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ20374.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000381; BAA91128.1; -.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20660 MW; 855494F0366B2291 CRC64;

Query Match
Best Local Similarity 39.4%; Score 339; DB 2; Length 199;
Matches 72; Conservative 26; Mismatches 56; Indels 18; Gaps 3;

QY 1 MEHYRKAGSVLPA-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
   ||:|||||
Db 1 MENFRKVRSEEPAGCGAEGGPGGFADLAPGAVHMRVKEGSKIRNLMAFATASMAQP 60
   ||:|||||

QY 51 SARHVFSGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQTEDSW--VPASPDGLDP- 107
   :||:|||||
Db 61 ATRAIVFGCGRVTTKVTCAEILKRRLAGLHQVTRLRYRSRVREVMQSLPPGPTQGQTPG 120

QY 108 -----LTVRRHVPVAVVLLSRDPLDPNECGYQPPGAPGLGSMPSSSCGPRS 154
   ||:|||||
Db 121 EPAASLSVLKNVPGLAILLSKDALDPRQPGYQPPNPHPGPSPPPAAPASKRS 172

RESULT 8
Q7PTH5
ID Q7PTH5 PRELIMINARY; PRT; 215 AA.
AC Q7PTH5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016002 (Fragment).
GN Name=ENSANG0000013513;
OS Anopheles gambiae str. PEST.
```

```
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008807; EAA03913.2; -.
FT NON_TER 1
FT NON_TER 215
SQ SEQUENCE 215 AA; 24004 MW; AA72B25FAA8ABFBE CRC64;

Query Match
Best Local Similarity 30.5%; Score 263; DB 2; Length 215;
Matches 65; Conservative 28; Mismatches 61; Indels 10; Gaps 5;

QY 1 MEHYRKAGSVLPA---PMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGSGARHVPF 57
   ||:|||||
Db 6 NMHYKKGKNVEEELSQEIQPIEVLPAFLMFLMHVKGSEVKNLVDYAKKALEEGTHRSVVW 65
   ||:|||||

QY 58 SSGSRAAGKAVSCAEIVKRRVPGHLQTLK---RFLQTEDSWVPASPDGLDPLTVRRHVP 115
   ||:|||||
Db 66 SSGDGGVGKTISCAEIMKRHFE-LHQVTRICYRKCVVEEFWDP--QQEGLEQIVAKRNIP 122
   ||:|||||

QY 116 AVVWLLSRDPLDPNECGYQPPGAPGL--GSMPSSSCGPRRRR 157
   ||:|||||
Db 123 CVHILMSLDEIDPSVAGYQHSKTQGGFWTGGGLSSDSDGPRKRTK 166
   ||:|||||

RESULT 9
Q9V9B6
ID Q9V9B6 PRELIMINARY; PRT; 206 AA.
AC Q9V9B6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG9422-PA (CG9422-pc).
GN ORFNames=CG9422;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```


RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF428441; AAL16210.1; -.
DR EMBL; AY054208; AAL06869.1; -.
DR EMBL; AF412102; AAL06555.1; -.
DR EMBL; AY124847; AAM70556.1; -.
DR HSSP; P74761; 1H0X.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 350 AA; 37383 MW; DB474865DF4E162E CRC64;

Query Match 23.9%; Score 206; DB 2; Length 350;
Best Local Similarity 29.9%; Pred. No. 1.3e-09;
Matches 53; Conservative 23; Mismatches 65; Indels 36; Gaps 4;

QY 1 MEHYRKAGSVLPAPSPMPQLPPDPTLEMRVRDGSKIRNLGLALGRLEGGSARHVVFGS 60
Db MDK YQR-----VVKPKADTPIDANEIRITTSQGRARNYITYAMTLLQDKSGSTEVEVFKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVL 120
Db GRAINKTVTIVELIKRRIPDLHQNTSIGSTDITDWEPT--EEGLLPLETTTRHVSMTIT 111

QY 121 LSRDPLDPNECGYQPP-----GAPPLGSMPSSSCGPRRRRAR 159
Db LSKIELNTSSVGVCPIPIELVKPMGIDIDYEGREGSPGGRG-----RGRGRGR 159

RESULT 15
Q7XJ13
ID Q7XJ13 PRELIMINARY; PRT; 732 AA.
AC Q7XJ13;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Chloroplast nucleoid DNA-binding protein-like protein.
GN Name=P0478E02.13;
OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Oliveira A.C., Mattos L.T., Carvalho F.F., Shimano A., Zimmer P.D.,
RA Malone G., Dellagostin O.;
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB109206; BAC79194.1; -.
DR Gramene; Q7XJ13; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR011574; Alba_DUF78.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp; 1.
DR ProDom; PD010497; Alba_DUF78; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW DNA-binding.
SQ SEQUENCE 732 AA; 78353 MW; 7EA86DD5386374AC CRC64;

Query Match 23.4%; Score 201.5; DB 2; Length 732;
Best Local Similarity 30.9%; Pred. No. 7.4e-09;
Matches 55; Conservative 19; Mismatches 79; Indels 25; Gaps 3;

QY 1 MEHYRKAGSVLPAPSPMPQLPPDPTLEMRVRDGSKIRNLGLALGRLEGGSARHVVFGS 60
Db MDY YQR-----VEKPREAPIKENEIRITTOGRMRNYITYATTLLQDKSGDEVVFKAM 543

QY 61 GRAAGKAVSCAEIVKRRVPGHQLTKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPAVWVL 120
Db GRAINKTVMIAELIKRRIVGLHQNTTGTSTDTITDWEPL--EEGLLPLETTTRHVSMTIT 601

QY 121 LSRDPLDPNECGYQPP-----GAPPLGSMPSSSCGPRRRRAR 162
Db LSKKELDTSSIGYQSPLPADKVKPLVEYENEEDAPSPAGRGRGGQGRGRGRGTR 659

Search completed: January 26, 2005, 15:06:25
Job time : 177.603 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:47 ; Search time 41.4407 Seconds
(without alignments)
260.851 Million cell updates/sec

Title: US-10-057-813-14
Perfect score: 861
Sequence: 1 MEHYRKAGSVLPAPSPMPQ.....SMPSSCGPRRRRRDTRS 163

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	642	74.6	128	4	US-09-621-976-4077 Sequence 4077, Ap
2	395	45.9	78	4	US-09-513-999C-7680 Sequence 7680, Ap
3	151.5	17.6	118	4	US-09-270-767-39818 Sequence 39818, A
4	151.5	17.6	118	4	US-09-270-767-55035 Sequence 55035, A
5	91.5	10.6	950	4	US-09-252-991A-25927 Sequence 25927, A
6	91	10.6	345	4	US-09-252-991A-18806 Sequence 18806, A
7	91	10.6	921	4	US-09-252-991A-17504 Sequence 17504, A
8	90.5	10.5	330	4	US-09-252-991A-19363 Sequence 19363, A
9	90	10.5	429	4	US-09-252-991A-20363 Sequence 20363, A
10	89.5	10.4	432	4	US-09-252-991A-23622 Sequence 23622, A
11	87	10.1	374	4	US-09-252-991A-28830 Sequence 28830, A
12	85.5	9.9	446	4	US-09-252-991A-27110 Sequence 27110, A
13	85.5	9.9	477	4	US-09-252-991A-19831 Sequence 19831, A
14	85.5	9.9	497	4	US-09-252-991A-23620 Sequence 23620, A
15	85	9.9	441	4	US-09-252-991A-28965 Sequence 28965, A
16	84.5	9.8	282	4	US-09-252-991A-19978 Sequence 19978, A
17	84	9.8	546	4	US-09-252-991A-25851 Sequence 25851, A
18	83.5	9.7	1482	4	US-09-410-551B-21 Sequence 21, Appl
19	83.5	9.7	1482	4	US-09-940-316B-21 Sequence 21, Appl
20	83.5	9.7	1488	4	US-09-410-551B-17 Sequence 17, Appl
21	83.5	9.7	1488	4	US-09-940-316B-17 Sequence 17, Appl
22	83.5	9.7	1509	4	US-09-410-551B-23 Sequence 23, Appl
23	83.5	9.7	1509	4	US-09-940-316B-23 Sequence 23, Appl
24	83.5	9.7	1517	4	US-09-410-551B-19 Sequence 19, Appl
25	83.5	9.7	1517	4	US-09-940-316B-19 Sequence 19, Appl
26	83.5	9.7	6396	4	US-09-410-551B-72 Sequence 72, Appl
27	83.5	9.7	6396	4	US-09-940-316B-72 Sequence 72, Appl

28	83	9.6	240	4	US-09-252-991A-31100 Sequence 31100, A
29	83	9.6	594	4	US-09-252-991A-32578 Sequence 32578, A
30	82.5	9.6	472	4	US-09-252-991A-16723 Sequence 16723, A
31	82	9.5	404	4	US-09-252-991A-29985 Sequence 29985, A
32	81.5	9.5	469	4	US-09-252-991A-32604 Sequence 32604, A
33	81.5	9.5	491	4	US-09-252-991A-31758 Sequence 31758, A
34	81	9.4	302	4	US-09-252-991A-25724 Sequence 25724, A
35	81	9.4	482	4	US-09-252-991A-29113 Sequence 29113, A
36	81	9.4	516	4	US-09-252-991A-19788 Sequence 19788, A
37	80.5	9.3	266	4	US-09-252-991A-30538 Sequence 30538, A
38	80.5	9.3	964	3	US-08-484-791-2 Sequence 2, Appli
39	80	9.3	313	4	US-09-252-991A-23631 Sequence 23631, A
40	80	9.3	750	4	US-09-252-991A-23762 Sequence 23762, A
41	80	9.3	860	4	US-09-252-991A-28607 Sequence 28607, A
42	79.5	9.2	251	4	US-09-252-991A-31800 Sequence 31800, A
43	79.5	9.2	377	4	US-09-252-991A-24675 Sequence 24675, A
44	79.5	9.2	400	4	US-09-252-991A-31900 Sequence 31900, A
45	79.5	9.2	416	4	US-09-252-991A-30923 Sequence 30923, A

ALIGNMENTS

RESULT 1

US-09-621-976-4077
; Sequence 4077, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4077
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52...-1
; NAME/KEY: UNSURE
; LOCATION: 72
; OTHER INFORMATION: Xaa = Asp,Glu,Gly,Val
US-09-621-976-4077

Query Match 74.6%; Score 642; DB 4; Length 128;
Best Local Similarity 99.2%; Pred. No. 4.5e-62;
Matches 124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MEHYRKAGSVLPAPSPMPQPPDPTLEMRVDRGSKIRNLGLALGRLEGGSARHVVFGS	60
Db	1	MEHYRKAGSVLPAPSPMPQPPDPTLEMRVDRGSKIRNLGLALGRLEGGSARHVVFGS	60
Qy	61	GRAAGKAVSCAEIVKRRVPGHLQTLKRLQTLQTESWVPASPDGLDPLTVRRHVPVWL	120
Db	61	GRAAGKAVSCAEIVKRRVPGHLQTLKRLQTLQTESWVPASPDGLDPLTVRRHVPVWL	120
Qy	121	LSRDP 125	
Db	121	LSRXP 125	

RESULT 2

US-09-513-999C-7680
; Sequence 7680, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.

```
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pn
; SEQ ID NO 7680
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7680

Query Match      45.9%; Score 395; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.6e-35;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60

QY 61 GRAAGKAVSCAEIVKRRV 78
Db 61 GRAAGKAVSCAEIVKRRV 78

RESULT 3
US-09-270-767-39818
; Sequence 39818, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39818
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39818

Query Match      17.6%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 7.6e-09;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS GRAAGKAVSC 70
Db 1 DLFFEDCMPKSKQDFLWMHVKGTKVSNVIEFAQEALNKGHRCVWWSGGSGGVGKTISC 60

QY 71 AEIVKRRVPGLHQLTKLRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 4
US-09-270-767-55035
; Sequence 55035, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55035
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55035

Query Match      17.6%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 7.6e-09;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS GRAAGKAVSC 70
Db 1 DLFFEDCMPKSKQDFLWMHVKGTKVSNVIEFAQEALNKGHRCVWWSGGSGGVGKTISC 60

QY 71 AEIVKRRVPGLHQLTKLRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 5
US-09-252-991A-25927
; Sequence 25927, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25927
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25927

Query Match      10.6%; Score 91.5; DB 4; Length 950;
Best Local Similarity 25.8%; Pred. No. 0.39;
Matches 40; Conservative 12; Mismatches 60; Indels 43; Gaps 6;

QY 15 PSPMPQLPPDTLEMRVRDGSKIRNLLGLALG--RLEGG SARHVVFGS GRAAGKAVSCAE 72
Db 338 PRPAPQLL-----YTRLRTALGAPAGMERLRGG---HCLQPATGAALGTEPLPA 384
QY 73 IVKRRV-----PGLHQLTKLRFQLQTEDSWVPASPDITGLDPLTVRRHVPVAVVLLSR- 123
Db 385 QVRRRSDEAAARQRPV-----ATGDPGPRPATQRRRTAPGLRTGRPOH 427
QY 124 --DPLDPNECGYQPPGAPPGGLGSMFSSSCGPRSR 156
Db 428 PGKPLRPGDAGHPQPGROPAPAEPELPLRPGKLR 462

RESULT 6
US-09-252-991A-18806
; Sequence 18806, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
```

```

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18806
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18806

Query Match      10.6%; Score 91; DB 4; Length 345;
Best Local Similarity 25.7%; Pred. No. 0.12;
Matches 45; Conservative 12; Mismatches 38; Indels 80; Gaps 8;

QY 4 YRKAGSVELPA-----PSPMP-----QL--PPDTLEMRVD-----GSKIRNLLGL 42
Db 228 HRGPGSQRLPAGRSRRGRRLVPQPVGAGQLLRPPAELPRLRGNGTEPAGTRSQGNPGK 287
QY 43 ALGRLEGGSARHVVFSGSGRAAGKAVSCAEIVKRRVPGHLQLTKLRFLOTEDSWVPASPD 102
Db 288 A-PELSGQGRRLRLRDAAGRRSGLA-----PATPA 316
QY 103 TGLDPLTVRRHVPVAVVLLSRDPLDPNECGYQPPGAPGLGSMPSSSCGPRSRRR 157
Db 317 AGLPAATGR-----GRRPGLGAVPLS---PAARRR 343

RESULT 7
US-09-252-991A-17504
; Sequence 17504, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17504
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: UNSURE
; LOCATION: (886)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-17504

Query Match      10.6%; Score 91; DB 4; Length 921;
Best Local Similarity 28.2%; Pred. No. 0.42;
Matches 57; Conservative 11; Mismatches 70; Indels 64; Gaps 11;

QY 11 ELPAPSPMPQL-----PPDTLEMRVRDGSKIRNLLGLALG-RLEGGSARHVVFSG 59
Db 164 DLPAPDRRPGAALQGPRPVQPPHPHRRQVR-----RRLHPLHAGLRQEGVHRRH----- 213
QY 60 SGRAAGKAVSCAEIVKR-----RVPGHLQLTKLRFLOTEDSWVPASPDITGL---DPLTVRR 112
Db 214 PGRPRRPLLRPAEPDHRRLARRPGLHRAALLRLQLRQHRHRRPDDGLPLHGELOPAL 273
QY 113 HVPA---VWVLLSRDPLD-----PNECGYQ-----PPGAPPG----- 141
Db 274 HQPVDHRVLAALAHQPVDLAARLPLHQPGRQPRQHLPLDLPQVPDHPAGRPVARRQLHLH 333
QY 142 -LGSM----PSSSCGPRSRRA 158
Db 334 HLGRLARHVAGDRAGPRSERRA 355
```

```

RESULT 8
US-09-252-991A-19363
; Sequence 19363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19363
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19363

Query Match      10.5%; Score 90.5; DB 4; Length 330;
Best Local Similarity 27.3%; Pred. No. 0.12;
Matches 44; Conservative 12; Mismatches 72; Indels 33; Gaps 7;

QY 5 RKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGSGRAA 64
Db 87 RAAGRLRIPRPAPDPRVAADA-ERRHRHGRPHPSGGAVLAALPSDVGGH---RPGRRAP 142
QY 65 GKAVSCAEIVKR---RVPGHL--QLTKLRFLOTEDSW--VPASPDITGLDPLTVRRHVPVAV 117
Db 143 GQAAPGKRRARRQRPRRPGRHAGSLGQRRSGIPEGAWRGPRPGPPAGRSPPAPR----- 197
QY 118 WVLLSRDPLDPNECGYQPPGAPGLGSMPSSSCGPRSRRA 158
Db 198 -----PAERRAAALRSAGQPP-----GDRSRRA 221
```

```

RESULT 9
US-09-252-991A-20363
; Sequence 20363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20363
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20363

Query Match      10.5%; Score 90; DB 4; Length 429;
Best Local Similarity 31.3%; Pred. No. 0.2;
Matches 51; Conservative 11; Mismatches 71; Indels 30; Gaps 8;

QY 12 LPA---PSPMPQLPPDTLEMRVRDGSKIR---NLLGLALGRLEGGSARHVVFSGS----G 61
Db 119 LPAGLRGGPVGALPAVDLHARRADLPRRRLRVLPQQAGQAAHLGQGVHRRFAGGDLPLPG 178
QY 62 RAAGKAVSCAEIVKRRVPGHLQLTKLRFLOTEDSWVPASPDITGLDPLTVRRHVPVAVVLL 121
Db 179 RGAGR-----LHRRHPGGEPLRRRRFAGL-DRAVPAVLPRPGPD-----RRLHPARLHLA 226
```


; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19831

Query Match 9.9%; Score 85.5; DB 4; Length 477;
Best Local Similarity 29.1%; Pred. No. 0.71;
Matches 39; Conservative 11; Mismatches 49; Indels 35; Gaps 7;

QY 43 ALGRLE-GGSARHVVFSGS-----RAAGKAVSCAEIVKRRVPGHLQHLTKLRLQTEDSW 96
Db ATGRLHPGGRRRRMARAVAGPREPVRPGAPAGAGGAGRRAPGLYQ----- 384
QY 97 VPASPDGLDP-----LTVRRHVPVAVVLLSRDPLDPNECGYQPPGAPPLGSMPS 148
Db -PCAAQRRGDPGRRPGAGRLGGQRRRPG-YLAPGREPGDPRR---TPAGALPGSGTAPG- 438
QY 149 SCGPRSRRRARDTR 162
Db 439 --GSHSRGRGRTPR 450

RESULT 14
US-09-252-991A-23620
; Sequence 23620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23620
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23620

Query Match 9.9%; Score 85.5; DB 4; Length 497;
Best Local Similarity 26.9%; Pred. No. 0.74;
Matches 49; Conservative 10; Mismatches 52; Indels 71; Gaps 11;

QY 19 POLPDDTLEMRV-----RDGSKIRNLLGLALGRLEG---GSARHVVFSGSGRAAGKAVSC 70
Db PQQPPGPLPARQGRPPGRAGSDLRRAFAATAGRSAGDRQGRRLL-----RAAGPG--- 293
QY 71 AEIVKRRVPGHLQHLTKLRLQTE-----DSW-----VP-----ASPDGLDP- 107
Db ---LPQRTPGLHARRQRRVAADPGPSVRAPAWRGGRHADLPGQPQARQLAEFGTGPAPA 350
QY 108 -----LTVRRHVPVAVVLLSRDPLDPNECGYQPPGAPPLGSMPS 156
Db RQSRLLRHLRLHRPA-----EGRGQHPRGA-----GRAPAMDAGHLRPRRR 394
QY 157 RA 158
Db 395 RA 396

RESULT 15
US-09-252-991A-28965
; Sequence 28965, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28965
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (338),(413)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-28965

Query Match 9.9%; Score 85; DB 4; Length 441;
Best Local Similarity 27.0%; Pred. No. 0.72;
Matches 47; Conservative 14; Mismatches 71; Indels 42; Gaps 8;
QY 5 RKAGSVELPAPSPMP-----QLPPDDTLEMRVRDGSKIRNLLGLALGRLEG 49
Db RRAGGAPAAEPGRLPRLRRRTARRRQPGGRQRTPEAAGGEGADGGRARQEPGQALGRRDP 209
QY 50 GSARHVVFSGSR-AAGKAVSCAEIVKRRVFGHLQHLTKLRLQTEDSWVPASPDGLDPL 108
Db 210 GTP---VGAGQGHPAAG-----GLRRLGQPRGLHRRRQHLPEP-----AH 247
QY 109 TVRRHVPVAVVLLSRDPLDPNECGYQPPGAPPLGSMPS 162
Db SAQRQEPAA--AAQSRSPRRSAAAG-QPRGTHRA-GIRPAVSEGRRLVARPRR 297

Search completed: January 26, 2005, 15:08:51
Job time : 42.4407 secs

This Page Blank (uspto)

C;Accession: S25185; S21012
R;Dammann, T.; Wohlleben, W.
Mol. Microbiol. 6, 2267-2278, 1992
A;Title: A metalloprotease gene from Streptomyces coelicolor "Mueller" and its transcrip
A;Reference number: S25185; MUID:93023855; PMID:1406267
A;Accession: S25185
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <DAM>
A;Cross-references: UNIPROT:P43167; EMBL:Z11929; NID:g46866; PIDN:CAA77983.1; PID:g46867
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992

Query Match 15.1%; Score 70; DB 2; Length 367;
Best Local Similarity 27.9%; Pred. No. 9.4;
Matches 31; Conservative 9; Mismatches 29; Indels 42; Gaps 6;

QY 7 AGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGL-----AL 44
Db 167 AARPQLPAP---PQLPPGLVPQLTPDRPLGRRALGLLHDLRHAAPRRRAGTLRGLAAL 223

QY 45 GRLEGG SARHVVFGSGRAAGKAV---SCAEIVK---RRV-----PGL 81
Db 224 PVLDDG---HALGRGEGTVAYRCAVGHGCEVLRLHLVRRVADEGCATRPGL 271

RESULT 7
T13713
beta3 protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13713
R;Zhang, Y.Q.; Broadie, K.S.
submitted to the EMBL Data Library, October 1998
A;Description: Dissecting the functions of beta3 of AP3 complex.
A;Reference number: Z17706
A;Accession: T13713
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-1160 <ZHA>
A;Cross-references: UNIPROT:O77290; EMBL:AJ011778; NID:e1330417; PID:e1330418; PIDN:CAAO
C;Genetics:
A;Gene: beta3
A;Cross-references: FlyBase:FBgn0003210
A;Introns: 214/3; 491/1; 899/1

Query Match 15.1%; Score 70; DB 2; Length 1160;
Best Local Similarity 33.8%; Pred. No. 32;
Matches 25; Conservative 6; Mismatches 19; Indels 24; Gaps 5;

QY 1 MEHYRK---AGSVLPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVV 57
Db 659 LSHYLNMPAAGYKELPA---FPVIPPD-----SSVRNIAGFMQDKLPGEDSP---- 702

QY 58 SG-----SGRAAGK 66
Db 703 SGHSKDKSGREGGK 716

RESULT 8
T10586
small nuclear ribonucleoprotein-associated protein homolog F9F13.90 - Arabidopsis thalia
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10586
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16991
A;Accession: T10586
A;Molecule type: DNA
A;Residues: 1-257 <BEV>
A;Cross-references: UNIPROT:Q9SUN5; EMBL:AL080253; GSPDB:GN00062; ATSP:F9F13.90
A;Experimental source: cultivar Columbia; BAC clone F9F13
C;Genetics:

A;Gene: ATSP:F9F13.90
A;Map position: 4
C;Superfamily: proline-rich protein

Query Match 15.0%; Score 69.5; DB 2; Length 257;
Best Local Similarity 29.6%; Pred. No. 7.3;
Matches 24; Conservative 10; Mismatches 26; Indels 21; Gaps 3;

QY 20 QLPPDTLEMRVRDGSKIRNLLGLAL-----GRLEGG SARHVVFGS 60
Db 50 KLPP-AKGKINEERDRRTGLVLRGEEVISMTVEGPPPEESRAKAGSAAVAGPGI 108

QY 61 GRAAGKAVSCAEIVKRRVPGL 81
Db 109 GRAAGRGVFTGPLVQAQ-PGL 128

RESULT 9
T02784
calcium-dependent protein kinase (EC 2.7.1.1-) - maize (strain W64A)
C;Species: Zea mays (maize)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T02784
R;Muriillo, I.; Jaeck, E.; Cordero, M.; San Segundo, B.
submitted to the EMBL Data Library, July 1998
A;Description: A calcium-dependent protein kinase possibly involved in pathogen defense
sis-related PRms gene.
A;Reference number: Z14736
A;Accession: T02784
A;Status: translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-639 <MUR>
A;Cross-references: UNIPROT:O82107; EMBL:AJ007366; PIDN:CAA07481.1
A;Experimental source: strain W64A; seed
C;Function:
A;Description: probably involved in pathogen defense in maize plants
C;Superfamily: calcium-dependent protein kinase; EF hand; phosphotransferase; serine/threonine-specifi.
C;Keywords: ATP; calcium binding; EF hand; phosphotransferase; serine/threonine-specifi.
F;151-411/Domain: protein kinase homology <KIN>
F;159-167/Region: protein kinase ATP-binding motif
F;454-486/Domain: calmodulin repeat homology <EF1>
F;490-522/Domain: calmodulin repeat homology <EF2>
F;526-558/Domain: calmodulin repeat homology <EF3>
F;560-592/Domain: calmodulin repeat homology <EF4>
F;182/Active site: Lys #status predicted

Query Match 14.9%; Score 69; DB 1; Length 639;
Best Local Similarity 25.6%; Pred. No. 22;
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVLPAPSPMPQLPPDTLEM-----RVRDGSKIRNLLGLALGRLEGGSA 52
Db 108 KPGGAANASPSRPRPQVKRVSSAGLLGSLVLRRTKNTENLKDYSGLRRLRGQGFQGT 167

QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78
Db 168 HLCV----ERATGKELACKSILKRKL 189

RESULT 10
T39026
conserved hypothetical protein SPAC6C3.02c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39026
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21750
A;Accession: T39026
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-172 <DEV>
A;Cross-references: UNIPROT:Q10307; EMBL:Z69731; PIDN:CAA93615.1; GSPDB:GN00066; SPDB:S

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:40 ; Search time 97.478 Seconds
(without alignments)
537.138 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPLHQLTKLRLQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	163	2 Q8N5L8	Q8n5l8 homo sapien
2	440	95.0	163	2 Q99JH1	Q99jhl mus musculu
3	439	94.8	163	2 Q9CYX4	Q9cyx4 mus musculu
4	264.5	57.1	224	2 Q6DGS1	Q6dgs1 brachydanio
5	224	48.4	199	2 Q91WE3	Q91we3 mus musculu
6	222	47.9	199	2 Q9BUL9	Q9bul9 homo sapien
7	218	47.1	199	2 Q9NX88	Q9nxx8 homo sapien
8	167	36.1	206	2 Q9V9B6	Q9v9b6 drosophila
9	166	35.9	215	2 Q7PTH5	Q7pth5 anopheles g
10	162	35.0	206	2 Q6NNC7	Q6nnc7 drosophila
11	162	35.0	206	2 AAR96155	Aar96155 drosophil
12	137.5	29.7	315	2 Q944A2	Q944a2 arabisdopsis
13	137.5	29.7	538	2 Q9LNU5	Q9lnu5 arabisdopsis
14	128.5	27.8	350	2 Q93VA8	Q93va8 arabisdopsis
15	125.5	27.1	485	2 Q9M8Z6	Q9m8z6 arabisdopsis
16	119.5	25.8	272	2 Q8H7C6	Q8h7c6 arabisdopsis
17	119.5	25.8	732	2 Q7XJ13	Q7xjl3 oryza sativ
18	111	24.0	241	2 Q7R6W6	Q7r6w6 plasmodium
19	111	24.0	248	2 Q8IAX8	Q8iax8 plasmodium
20	109.5	23.7	262	2 Q7XT90	Q7xt90 oryza sativ
21	94	20.3	211	2 Q8IDN4	Q8idn4 plasmodium
22	90	19.4	202	2 Q7RPX4	Q7rpx4 plasmodium
23	86.5	18.7	369	2 Q9LQR7	Q9lqr7 arabisdopsis
24	86	18.6	160	2 Q9DEP6	Q9dep6 sparus aura
25	83	17.9	534	2 Q9ZPM0	Q9zpm0 mesembryant
26	82	17.7	275	2 Q9AWP6	Q9awp6 oryza sativ
27	77.5	16.7	206	2 Q7YWL9	Q7ywl9 caenorhabdi
28	74	16.0	296	2 Q82JH1	Q82jhl streptomyce
29	74	16.0	321	2 Q6YYK9	Q6yyk9 oryza sativ
30	74	16.0	321	2 BAD05714	Bad05714 oryza sat
31	74	16.0	321	2 BAD05796	Bad05796 oryza sat

32	73.5	15.9	249	1	GPPI_YEAST	P41277 saccharomyc
33	73	15.8	907	2	Q9C946	Q9c946 arabidopsis
34	72	15.6	453	2	Q82CD4	Q82cd4 streptomyce
35	72	15.6	458	2	Q72DN2	Q72dn2 desulfovibr
36	72	15.6	458	2	AAS95377	Aas95377 desulfovi
37	72	15.6	521	2	Q94KH6	Q94kh6 lycopersico
38	71.5	15.4	110	2	Q9Y965	Q9y965 aeropyrum p
39	71.5	15.4	249	2	Q8LN54	Q8ln54 oryza sativ
40	71.5	15.4	812	2	Q9FBT4	Q9fbt4 streptomyce
41	71	15.3	295	2	Q72B31	Q72b31 desulfovibr
42	71	15.3	295	2	AAS96284	Aas96284 desulfovi
43	70.5	15.2	216	2	Q6ZHB3	Q6zhb3 oryza sativ
44	70.5	15.2	216	2	BAD07586	Bad07586 oryza sat
45	70.5	15.2	251	2	Q7VYE7	Q7vye7 bordetella

ALIGNMENTS

RESULT 1

Q8N5L8	PRELIMINARY;	PRT;	163 AA.
ID Q8N5L8			
AC Q8N5L8;			
DT 01-OCT-2002	(Tremblrel. 22, Created)		
DT 01-OCT-2002	(Tremblrel. 22, Last sequence update)		
DT 01-JUN-2003	(Tremblrel. 24, Last annotation update)		
DE Chromosome 9	open reading frame 23 protein.		
GN Name=C9orf23;			
OS Homo sapiens (Human)			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Skin;			
RX MEDLINE=22388257; PubMed=12477932;			
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,			
RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA Jones S.J., Marra M.A.;			
RT "Generation and initial analysis of more than 15,000 full-length human			
RT and mouse cDNA sequences."			
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Skin;			
RA Strausberg R.;			
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BC032136; AAH32136.1; -.			
DR Genew; HGNC:19909; C9orf23.			
DR InterPro; IPR011574; Alba DUF78.			
DR ProDom; PD010497; Alba DUF78; 1.			
SQ SEQUENCE 163 AA; 17631 MW; 442C8727191A0BCE CRC64;			

Query Match 100.0%; Score 463; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1e-39;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPTLEMRVDRDGSKIRNLLGLALGRLEGCSARHVVFGS 60
|||||

Db 1 MEHYRKAGSVLPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
|||||

Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
|||||

RESULT 2

Q99JH1 PRELIMINARY; PRT; 163 AA.

AC Q99JH1;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein (RIKEN cDNA 2810432D09).

GN Name=2810432D09Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Skeletal muscle;

RA Ievolella C., Zara I., Lanfranchi G.;

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.;

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.;"

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.;

RA Strausberg R.;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ293897; CAC34588.1; -

DR EMBL; BC047068; AAH47068.1; -

DR MGD; MGI:1917211; 2810432D09Rik.

DR InterPro; IPR011574; Alba_DUF78.

DR ProDom; PD010497; Alba_DUF78; 1.

KW Hypothetical protein.

SQ SEQUENCE 163 AA; 17675 MW; 87675201AF87F5B6 CRC64;

Query Match 95.0%; Score 440; DB 2; Length 163;

Best Local Similarity 95.6%; Pred. No. 2.4e-37;

Matches 87; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVLPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
|||

Db 1 MEQYRRAGSVLPASSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSSTRHVVFSGS 60
|||

QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
|||||

Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
|||||

RESULT 3

Q9CYX4 PRELIMINARY; PRT; 163 AA.

AC Q9CYX4;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length

DE enriched library, clone:2810432D09 product:hypothetical protein, full

DE insert sequence.

GN Name=2810432D09Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.;"

RL Meth. Enzymol. 303:19-44 (1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.;"

RL Nature 409:685-690 (2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA The FANTOM Consortium,

RT the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.;"

RL Nature 420:563-573 (2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Muramatsu M., Hayashizaki Y.;

RA Konno H., Okazaki Y., Sugahara Y., Shibata K., Itoh M.,

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.;"

RL Genome Res. 10:1617-1630 (2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nishikawa S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.;"

RL Genome Res. 10:1757-1771 (2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013222; BAB28723.1; -;
DR MGD; MGI:1917211; 2810432D09Rik.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 163 AA; 17648 MW; DC89421F4F617429 CRC64;

Query Match 94.8%; Score 439; DB 2; Length 163;
Best Local Similarity 95.6%; Pred. No. 3e-37; 3; Indels 0; Gaps 0;
Matches 87; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALRLEGSSARHVVVFGS 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MEQYREAGSVELPASSPMPQLPPDTLEMRVRDGSKIRNLLGLALRLEGSSRHHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91

RESULT 4
Q6DGS1
ID Q6DGS1 PRELIMINARY; PRT; 224 AA.
AC Q6DGS1;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC076268; AAH76268.1; -.
KW Hypothetical protein.
SQ SEQUENCE 224 AA; 24369 MW; C2FD38C9D8B2F38A CRC64;

Query Match 57.1%; Score 264.5; DB 2; Length 224;
Best Local Similarity 48.1%; Pred. No. 4.3e-19;
Matches 52; Conservative 17; Mismatches 20; Indels 19; Gaps 1;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALRLEGSSA----- 52
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MENYRKANTIEQPCPCFPDLPSTDPVVRKDGSKIRNLMRFALSRMEETAASADHEGS 60
QY 53 -----RHVVFGSGRAAGKAVSCAEIVKRRVPGHLQTLKRLF 89
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 EVSVGTGDNLCRQIVFTGVGQSVAKAITCVEIMKRRIRHGLHQLTKLAY 108

RESULT 5
Q91WE3
ID Q91WE3 PRELIMINARY; PRT; 199 AA.
AC Q91WE3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribonuclease P 25kDa subunit (Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone:C330033C13 product:hypothetical protein, full insert sequence).
DE length enriched library, clone:C330033C13 product:hypothetical protein, full insert sequence).
GN Name=Rpp25;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA the FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016085; AAH16085.1; -.
DR EMBL; AK082822; BAC38636.1; -.
DR MGD; MGI:2143151; AI851155.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21037 MW; B9DE105BE3293D64 CRC64;

Query Match 48.4%; Score 224; DB 2; Length 199;
Best Local Similarity 47.5%; Pred. No. 5.7e-15;
Matches 47; Conservative 14; Mismatches 28; Indels 10; Gaps 1;
QY 1 MEHYRKAGSVELPA-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
Db 1 MENFRKVRSEAPAGDGDGEGSPNSGPFADLPAGVAMRVKESKIRNLLAFATASMAQP 60
QY 51 SARHVVFSGSRAAGKAVSCAEIVKRRVPGHLQTLKRLF 89
Db 61 ATRAIVFSGCGRATTKVTTCAEILKRRLAGLHQVTRLRY 99

RESULT 6
Q9BUL9
ID Q9BUL9 PRELIMINARY; PRT; 199 AA.
AC Q9BUL9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE RPP25 protein (RNase P protein subunit p25).
GN Name=RPP25;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21998191; PubMed=12003489;
RA Guerrier-Takada C., Eder P.S., Gopalan V., Altman S.;
RT "Purification and characterization of Rpp25, an RNA-binding protein
RT subunit of human ribonuclease P.";
RL RNA 8:290-295(2002).
DR EMBL; BC002497; AAH02497.1; -.
DR EMBL; BC007270; AAH07270.1; -.
DR EMBL; AY034074; AAK54443.1; -.
DR IntAct; Q9BUL9; -.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
SQ SEQUENCE 199 AA; 20632 MW; 9A4494F0297B2A81 CRC64;

Query Match 47.9%; Score 222; DB 2; Length 199;
Best Local Similarity 46.5%; Pred. No. 9.1e-15;
Matches 46; Conservative 15; Mismatches 28; Indels 10; Gaps 1;
QY 1 MEHYRKAGSVELPA-----PSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG 50
Db 1 MENFRKVRSEAPAGCGAEGGPGSGPFADLPAGVAMRVKESKIRNLMFATASMAQP 60
QY 51 SARHVVFSGSRAAGKAVSCAEIVKRRVPGHLQTLKRLF 89
Db 61 ATRAIVFSGCGRATTKVTTCAEILKRRLAGLHQVTRLRY 99

RESULT 7
Q9NX88
ID Q9NX88 PRELIMINARY; PRT; 199 AA.
AC Q9NX88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)


```
RESULT 9
Q7PTH5
ID Q7PTH5 PRELIMINARY; PRT; 215 AA.
AC Q7PTH5;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ENSANGP0000016002 (Fragment).
GN Name=ENSANGG0000013513;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008807; EAA03913.2; -.
FT NON_TER 1
FT NON_TER 215
SQ SEQUENCE 215 AA; 24004 MW; AA72B25FAA8ABFBE CRC64;

Query Match 35.9%; Score 166; DB 2; Length 215;
Best Local Similarity 43.3%; Pred. No. 5.9e-09;
Matches 39; Conservative 16; Mismatches 31; Indels 4; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---PMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVV 57
Db 6 MMHYKKGKNEEELSQEQIPIEVLFPANFLWMHVKGSEVKNLVDYAKKALEEGTHRSVV 65

QY 58 SGSGRAAGKAVSCAEIVKRRVPGHLQTLKL 87
Db 66 SGSDGGVGKTTISCAEIMKRHFE-LHQVTRI 94

RESULT 10
Q6NNC7
ID Q6NNC7 PRELIMINARY; PRT; 206 AA.
AC Q6NNC7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE RE65722p.
GN Name=CG9422;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011363; AAR96155.1; -.
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;

Query Match 35.0%; Score 162; DB 2; Length 206;
Best Local Similarity 40.6%; Pred. No. 1.5e-08;
Matches 39; Conservative 16; Mismatches 33; Indels 8; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---LPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSAR 53
Db 1 MMHYRKAENVEKELSKSDLPEFDCMPKSKQDFLWMHVKGTKVSNVIEFAQEALNKGEHR 60

QY 54 HVVFGSGGRAAGKAVSCAEIVKRRVPGHLQTLKRF 89
Db 61 CVVWSGGGGVVKTTICCAEVLKRSHPLYQVTRMAY 95

RESULT 11
AAR96155
ID AAR96155 PRELIMINARY; PRT; 206 AA.
AC AAR96155;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE RE65722p.
GN CG9422.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011363; AAR96155.1; -.
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;

Query Match 35.0%; Score 162; DB 2; Length 206;
Best Local Similarity 40.6%; Pred. No. 1.5e-08;
Matches 39; Conservative 16; Mismatches 33; Indels 8; Gaps 2;

QY 1 MEHYRKAGSVLPAPS---LPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSAR 53
Db 1 MMHYRKAENVEKELSKSDLPEFDCMPKSKQDFLWMHVKGTKVSNVIEFAQEALNKGEHR 60

QY 54 HVVFGSGGRAAGKAVSCAEIVKRRVPGHLQTLKRF 89
Db 61 CVVWSGGGGVVKTTICCAEVLKRSHPLYQVTRMAY 95

RESULT 12
Q944A2
ID Q944A2 PRELIMINARY; PRT; 315 AA.
AC Q944A2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE At1g20220/T20H2 3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439833; AAL27504.1; -.
DR HSSP; O28323; INFJ.
DR InterPro; IPR011574; Alba DUF78.
DR ProDom; PD010497; Alba DUF78; 1.
SQ SEQUENCE 315 AA; 33763 MW; 6DC486E51238A937 CRC64;

Query Match 29.7%; Score 137.5; DB 2; Length 315;
Best Local Similarity 34.1%; Pred. No. 7.7e-06;
Matches 30; Conservative 21; Mismatches 28; Indels 9; Gaps 3;

QY 1 MEHYRKAGSVLP-APSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSG 59
```


Db 1 MDKYQR---VEKPKADTPIAE-----NEIRITSMGRARNYITYAMALLQENKSNEVIFKA 52

QY 60 SGRAAGKAVSCAEIVKRRVPGHLQLTKL 87

Db 53 MGRAINKSVTIVELIKRRIPGLHQITSI 80

RESULT 13

Q9LNU5

ID Q9LNU5 PRELIMINARY; PRT; 538 AA.

AC Q9LNU5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE T20H2.2 protein.

GN Name=T20H2.2;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S., Pham P.,

RA Toriumi M., Yu G., Chin C., Chiou J., Choi E., Chung M., Gonzalez A.,

RA Howng B., Liu A., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,

RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,

RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,

RA Ecker J.R., Federspiel N.A., Theologis A.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC022472; AAF79893.1; -.

DR PIR; H86335; H86335.

DR HSSP; O28323; 1NEJ.

DR InterPro; IPR011574; Alba DUF78.

DR ProDom; PD010497; Alba DUF78; 1.

SQ SEQUENCE 538 AA; 58572 MW; 1B0BB141F0F9AF22 CRC64;

Query Match 29.7%; Score 137.5; DB 2; Length 538;

Best Local Similarity 34.1%; Pred. No. 1.3e-05;

Matches 30; Conservative 21; Mismatches 28; Indels 9; Gaps 3;

QY 1 MEHYRKAGSVELP-APSENPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSG 59

Db 224 MDKYQR---VEKPKADTPIAE-----NEIRITSMGRARNYITYAMALLQENKSNEVIFKA 275

QY 60 SGRAAGKAVSCAEIVKRRVPGHLQLTKL 87

Db 276 MGRAINKSVTIVELIKRRIPGLHQITSI 303

RESULT 14

Q93VA8

ID Q93VA8 PRELIMINARY; PRT; 350 AA.

AC Q93VA8;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Atlg76010/T4012.22.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,

RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,

RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,

RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,

RA Ecker J.R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,

RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,

RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,

RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,

RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,

RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,

RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,

RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,

RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,

RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,

RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,

RA Ecker J.R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,

RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,

RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,

RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,

RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,

RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,

RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF428441; AAL16210.1; -.

DR EMBL; AY054208; AAL06869.1; -.

DR EMBL; AF412102; AAL06555.1; -.

DR EMBL; AY124847; AAM70556.1; -.

DR HSSP; P74761; 1HOX.

DR InterPro; IPR011574; Alba DUF78.

DR ProDom; PD010497; Alba DUF78; 1.

SQ SEQUENCE 350 AA; 37383 MW; DB474865DF4E162E CRC64;

Query Match 27.8%; Score 128.5; DB 2; Length 350;

Best Local Similarity 32.2%; Pred. No. 7.3e-05;

Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFSGS 60

Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVFKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQLTKL 87

Db 54 GRAINKTIVTIVELIKRRIPDLHQNTSI 80

RESULT 15

Q9M8Z6

ID Q9M8Z6 PRELIMINARY; PRT; 485 AA.

AC Q9M8Z6;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE F17A9.19 protein.

GN Name=F17A9.19;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B., Wu D.,

Search completed: January 26, 2005, 15:06:27
Job time : 99.478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 90.6915 Seconds
(without alignments)
359.949 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGHLQTLKRLFLQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	463	100.0	163	4	AAG89339	Aag89339 Human sec
2	463	100.0	163	6	ABU00014	Abu00014 Human nov
3	463	100.0	219	8	ADM90988	Adm90988 Human pha
4	457	98.7	163	2	AAY36104	Aay36104 Extended
5	457	98.7	163	8	ADP19412	Adp19412 Human sec
6	402	86.8	79	2	AAY11950	Aay11950 Human 5'
7	395	85.3	78	3	AAG03599	Aag03599 Human sec
8	218	47.1	211	4	AAM38973	Aam38973 Human pol
9	164	35.4	205	4	ABB64410	Abb64410 Drosophil
10	134.5	29.0	99	3	AAG36977	Aag36977 Arabidops
11	128.5	27.8	350	3	AAG16838	Aag16838 Arabidops
12	128.5	27.8	350	3	AAG51875	Aag51875 Arabidops
13	128.5	27.8	350	3	AAG51847	Aag51847 Arabidops
14	128.5	27.8	350	8	ADN72883	Adn72883 Thale cre
15	128.5	27.8	350	8	ADN72183	Adn72183 Thale cre
16	128.5	27.8	356	3	AAG51846	Aag51846 Arabidops
17	119	25.7	74	3	AAG36978	Aag36978 Arabidops
18	116	25.1	267	3	AAG33136	Aag33136 Zea mays
19	104.5	22.6	259	3	AAG34072	Aag34072 Zea mays
20	104.5	22.6	313	3	AAG34071	Aag34071 Zea mays
21	104	22.5	239	3	AAG33137	Aag33137 Zea mays
22	94	20.3	314	3	AAG51848	Aag51848 Arabidops
23	94	20.3	314	3	AAG16839	Aag16839 Arabidops
24	94	20.3	314	3	AAG51876	Aag51876 Arabidops
25	78	16.8	298	3	AAG51877	Aag51877 Arabidops

26	78	16.8	298	3	AAG16840	Aag16840 Arabidops
27	76	16.4	247	2	AAY48525	Aay48525 Human bre
28	75.5	16.3	255	6	ABU32060	Abu32060 Protein e
29	75.5	16.3	261	7	ABO65891	Abos65891 Klebsiell
30	75.5	16.3	494	5	ABP62768	Abp62768 Protein f
31	75.5	16.3	494	7	ADJ72179	Adj72179 Streptomy
32	73.5	15.9	121	6	ABU00334	Abu00334 Human nov
33	73.5	15.9	271	2	AAW57327	Aaw57327 Mitochond
34	73.5	15.9	271	2	AAW30683	Aaw30683 Glycerol-
35	73.5	15.9	271	2	AAW60263	Aaw60263 Klebsiell
36	73.5	15.9	271	2	AAW26169	Aay26169 Cytosolic
37	72.5	15.7	290	7	ABO74266	Abos74266 Pseudomon
38	72	15.6	164	3	AAW44836	Aay44836 Human Per
39	72	15.6	213	6	ADA54493	Ada54493 Human pro
40	72	15.6	215	3	AAG33138	Aag33138 Zea mays
41	71	15.3	240	5	ADK35009	Adk35009 Novel hum
42	71	15.3	566	7	ABO69785	Abos69785 Pseudomon
43	70.5	15.2	655	4	ABG23459	Abg23459 Novel hum
44	70.5	15.2	1203	7	ADJ70368	Adj70368 Human hea
45	70	15.1	246	3	AAG48477	Aag48477 Arabidops

ALIGNMENTS

RESULT 1
AAG89339
ID AAG89339 standard; protein; 163 AA.
XX
AC AAG89339;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 459.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
(GEST) GENSET.
Dumas Milne Edwards J, Bougueleret L, Jobert S;
WPI; 2001-367870/38.
N-PSDB; AAH64942.
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Claim 21; Page 910-911; 921pp; English.
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention
XX
SQ Sequence 163 AA;

Query Match 100.0%; Score 463; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 7.6e-47;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91

RESULT 2
ABU00014
ID ABU00014 standard; protein; 163 AA.
XX
AC ABU00014;
XX
DT 17-JAN-2003 (first entry)
XX
DE Human novel polypeptide #107.
XX
KW Human; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.
XX
OS Homo sapiens.
XX
PN WO200274961-A1.
XX
PD 26-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US005109.
XX
PR 15-MAR-2001; 2001US-00810173.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2003-040556/03.
DR N-PSDB; ABX05092.
XX
PT New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
PS Claim 9; SEQ ID NO 633; 235pp; English.
XX
CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative

CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABG99888-ABG99989 and ABU0010-ABU00433 represent human
CC polypeptides of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied by the European Patent Office
XX
SQ Sequence 163 AA;

Query Match 100.0%; Score 463; DB 6; Length 163;
Best Local Similarity 100.0%; Pred. No. 7.6e-47;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91
Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFQ 91

RESULT 3
ADM90988
ID ADM90988 standard; protein; 219 AA.
XX
AC ADM90988;
XX
DT 03-JUN-2004 (first entry)
XX

DE Human pharmaceutically useful protein SeqID 381.
XX
KW human; cancer; haematopoiesis; thrombosis; anaemia;
KW cardiovascular disorder; ischaemic heart disease;
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;
KW cystic fibrosis; chronic renal failure; glomerulopathy;
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
KW HIV infection; systemic lupus erythematosus; endocrine system;
KW diabetes mellitus; epilepsy; Alzheimer's disease;
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
KW fungal; parasitic; viral infection; cytostatic; anticoagulant;
KW thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;
KW immunosuppressive; antiallergic; dermatological; antirheumatic;
KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;
KW antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;
KW gene therapy; vaccine.
XX
OS Homo sapiens.

XX
PN WO2004020595-A2.

XX
PD 11-MAR-2004.

XX
PF 28-AUG-2003; 2003WO-US027107.

XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410951P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411111P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Hestir K;
XX
DR WPI; 2004-257410/24.
DR N-PSDB; ADM90779, ADM91197.
XX
XX New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 381; 254pp; English.
XX
CC This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. chronic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC antianaemic, cardiant, vasotropic, antiasthmatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polypeptide is a human protein of the invention.
XX
SQ Sequence 219 AA;

Query Match 100.0%; Score 463; DB 8; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.1e-46;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
57 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 116

QY 61 GRAAGKAVSCAEIVKRRVPGLHQLTKLRFQ 91
Db ||||||||||||||||||||||||||||||||||||||||||||
117 GRAAGKAVSCAEIVKRRVPGLHQLTKLRFQ 147

RESULT 4
AAY36104
ID AAY36104 standard; protein; 163 AA.
XX
AC AAY36104;
XX
DT 13-SEP-1999 (first entry)
XX

DE Extended human secreted protein sequence, SEQ ID NO. 489.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease.
XX
OS Homo sapiens.
XX
PN WO9931236-A2.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB002122.
XX
PR 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
DR N-PSDB; AAX97788.
XX
PT New isolated human secreted proteins.
XX
PS Claim 9; Page 423-424; 516pp; English.
XX
CC This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases
XX
SQ Sequence 163 AA;

Query Match 98.7%; Score 457; DB 2; Length 163;
Best Local Similarity 98.9%; Pred. No. 3.9e-46;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGLHQLTKLRFQ 91
Db ||||||||||||||||||||||||||||||||||||||||||||
61 GRAAGKAVSCAEIVKRRVPGLHQLTKLRFQ 91

RESULT 5
ADP19412
ID ADP19412 standard; protein; 163 AA.
XX
AC ADP19412;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human secreted polypeptide #263.
XX
KW Human; secreted protein; genetic disease.
XX

OS Homo sapiens.
XX US2004110939-A1.
PN 10-JUN-2004.
XX 15-OCT-2001; 2001US-00978360.
PF 17-DEC-1998; 98WO-IB002122.
XX 09-FEB-1999; 99WO-IB000282.
PR 21-JUN-2000; 2000WO-IB000951.
XX 15-SEP-2000; 2000US-00663600.
PA (GEST) GENSET SA.
XX Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;
PI Duclert A;
XX WPI; 2004-440404/41.
DR N-PSDB; ADP19007.
XX New isolated polynucleotide encoding secreted polypeptide, useful for
PT gene therapy, or in diagnostic procedures to identify individuals having
PT genetic diseases resulting from abnormal expression of the genes.
XX Claim 2; SEQ ID NO 668; 113pp; English.
PS The invention relates to human cDNA sequences that encode human secreted
XX proteins. The invention also relates to an antibody that specifically
CC binds to a polypeptide of the invention and a method of binding the
CC polypeptide to an antibody. The polynucleotides are useful for expressing
CC the entire secreted proteins which they encode and for distinguishing
CC human tissues and cells from non-human tissues and cells, and for
CC distinguishing between human tissues and cells that do or do not express
CC the polynucleotides comprising the cDNAs. The polynucleotides and
CC polypeptides are useful in forensic procedures or diagnostic procedures
CC to identify individuals with genetic diseases resulting from abnormal
CC expression of the genes corresponding to the cDNAs. The sequences are
CC also useful in gene therapy to control or treat genetic diseases. This
CC sequence represents a human secreted polypeptide of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 163 AA;
Query Match 98.7%; Score 457; DB 8; Length 163;
Best Local Similarity 98.9%; Pred. No. 3.9e-46;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLO 91
Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLO 91
RESULT 6
AAV11950
ID AAV11950 standard; protein; 79 AA.
XX AAV11950;
AC AAV11950;
XX 18-JUN-1999 (first entry)
DT Human 5' EST secreted protein SEQ ID No: 550.
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX Homo sapiens.
OS WO9906550-A2.
XX 11-FEB-1999.
PD 31-JUL-1998; 98WO-IB001232.
XX 01-AUG-1997; 97US-00905144.
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Duclert A, Lacroix B;
XX WPI; 1999-153780/13.
DR N-PSDB; AAX40672.
XX New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
XX Claim 34; Page 648; 675pp; English.
PS AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 79 AA;
Query Match 86.8%; Score 402; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 5.9e-40;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVP 79
Db 61 GRAAGKAVSCAEIVKRRVP 79
RESULT 7
AAG03599
ID AAG03599 standard; protein; 78 AA.
XX AAG03599;
AC AAG03599;
XX 06-OCT-2000 (first entry)
DT Human secreted protein, SEQ ID NO: 7680.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX Homo sapiens.
OS

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-01322484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.8%; Score 128.5; DB 3; Length 350;
Best Local Similarity 32.2%; Pred. No. 1.4e-06;
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;
QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVVFKAM 53
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLK 87
Db 54 GRAINKTIVTIVELIKRRIPDLHQNTSI 80

RESULT 14
ADN72883
ID ADN72883 standard; protein; 350 AA.
XX
AC ADN72883;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 778.
XX
KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
KW animal feed product; thale cress; cell wall biosynthesis;
KW nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
PN WO2004035798-A2.
XX
PD 29-APR-2004.
XX
PF 20-OCT-2003; 2003WO-EP011658.
XX
PR 18-OCT-2002; 2002EP-00079408.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Inze D, De Veylder L, Vlieghe K;
XX
DR WPI; 2004-348466/32.
XX
DR N-PSDB; ADN72882.
XX
PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
PS Claim 1; SEQ ID NO 778; 134pp; English.

XX
CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
CC the E2Fa/Dpa transcription factor, given in an exemplification of the
CC invention.
XX
SQ Sequence 350 AA;

Query Match 27.8%; Score 128.5; DB 8; Length 350;
Best Local Similarity 32.2%; Pred. No. 1.4e-06;
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;
QY 1 MEHYRKAGSVELPAPSPMPQLPDDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVVFKAM 53
QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLK 87
Db 54 GRAINKTIVTIVELIKRRIPDLHQNTSI 80

RESULT 15
ADN72183
ID ADN72183 standard; protein; 350 AA.
XX
AC ADN72183;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 78.
XX
KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
KW animal feed product; thale cress; cell wall biosynthesis;
KW nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
PN WO2004035798-A2.
XX
PD 29-APR-2004.
XX
PF 20-OCT-2003; 2003WO-EP011658.
XX
PR 18-OCT-2002; 2002EP-00079408.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Inze D, De Veylder L, Vlieghe K;
XX
DR WPI; 2004-348466/32.
XX
DR N-PSDB; ADN72182.
XX
PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
XX

```
PS Claim 1; SEQ ID NO 78; 134pp; English.
XX
CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
CC the E2Fa/Dpa transcription factor, given in an exemplification of the
CC invention.
XX
SQ Sequence 350 AA;

Query Match      27.8%; Score 128.5; DB 8; Length 350;
Best Local Similarity 32.2%; Pred. No. 1.4e-06;
Matches 28; Conservative 15; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
   |::: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MDKYQR-----VVKPKADTPIDANEIRITSQGRARNYITYAMTLLQDKGSTEVVEFKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLTKL 87
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54 GRAINKTVTIVELIXRRIPDLHQNTSI 80

Search completed: January 26, 2005, 15:01:05
Job time : 92.6915 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:06:38 ; Search time 75.8847 Seconds
(without alignments)
433.254 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGLHQLTKLRFQ 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	463	100.0	163	9	US-09-731-872-459	Sequence 459, App
2	463	100.0	163	10	US-09-876-997-459	Sequence 459, App
3	457	98.7	163	11	US-09-978-360A-668	Sequence 668, App
4	305	65.9	72	17	US-10-425-115-237341	Sequence 237341,
5	134.5	29.0	281	15	US-10-424-599-237821	Sequence 237821,
6	124	26.8	671	16	US-10-437-963-109378	Sequence 109378,
7	123	26.6	359	17	US-10-425-115-231292	Sequence 231292,
8	119.5	25.8	269	16	US-10-437-963-156973	Sequence 156973,
9	117.5	25.4	242	17	US-10-425-115-254668	Sequence 254668,
10	116.5	25.2	245	15	US-10-424-599-229595	Sequence 229595,
11	116.5	25.2	263	16	US-10-437-963-152741	Sequence 152741,
12	116	25.1	261	16	US-10-437-963-184232	Sequence 184232,
13	116	25.1	265	17	US-10-425-115-201476	Sequence 201476,

14	116	25.1	294	15	US-10-425-114-60641	Sequence 60641, A
15	114.5	24.7	178	17	US-10-425-115-232929	Sequence 232929, A
16	112	24.2	320	17	US-10-739-930-10929	Sequence 10929, A
17	109.5	23.7	262	16	US-10-437-963-191078	Sequence 191078, A
18	107	23.1	261	16	US-10-767-701-42504	Sequence 42504, A
19	105.5	22.8	276	17	US-10-425-115-308895	Sequence 308895, A
20	105.5	22.8	276	17	US-10-425-115-308929	Sequence 308929, A
21	105.5	22.8	278	15	US-10-425-114-60670	Sequence 60670, A
22	105.5	22.8	301	15	US-10-425-114-61107	Sequence 61107, A
23	105.5	22.8	315	15	US-10-425-114-54515	Sequence 54515, A
24	105.5	22.8	315	15	US-10-425-114-58679	Sequence 58679, A
25	105.5	22.8	325	15	US-10-425-114-60913	Sequence 60913, A
26	104	22.5	287	16	US-10-437-963-152743	Sequence 152743, A
27	103.5	22.4	275	16	US-10-767-701-45240	Sequence 45240, A
28	100	21.6	318	17	US-10-425-115-266917	Sequence 266917, A
29	99.5	21.5	261	17	US-10-425-115-308917	Sequence 308917, A
30	82.5	17.8	408	16	US-10-437-963-117756	Sequence 117756, A
31	82	17.7	140	16	US-10-437-963-185996	Sequence 185996, A
32	82	17.7	198	17	US-10-425-115-365780	Sequence 365780, A
33	75.5	16.3	255	15	US-10-282-122A-59984	Sequence 59984, A
34	74.5	16.1	137	17	US-10-425-115-308595	Sequence 308595, A
35	74.5	16.1	244	16	US-10-767-701-56157	Sequence 56157, A
36	74	16.0	247	15	US-10-425-114-48205	Sequence 48205, A
37	74	16.0	296	14	US-10-156-761-10321	Sequence 10321, A
38	73.5	15.9	126	17	US-10-425-115-268190	Sequence 268190, A
39	73.5	15.9	271	10	US-09-308-207-33	Sequence 33, Appl
40	73.5	15.9	271	14	US-10-369-493-22102	Sequence 22102, A
41	73	15.8	149	16	US-10-437-963-167180	Sequence 167180, A
42	72	15.6	213	14	US-10-094-749-2061	Sequence 2061, Ap
43	72	15.6	453	14	US-10-156-761-12950	Sequence 12950, A
44	71.5	15.4	559	14	US-10-369-493-19468	Sequence 19468, A
45	70.5	15.2	1203	16	US-10-408-765A-2174	Sequence 2174, Ap

ALIGNMENTS

RESULT 1

US-09-731-872-459
; Sequence 459, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-731-872-459

Query Match 100.0%; Score 463; DB 9; Length 163;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGSARHVVFGS 60
|||
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGSARHVVFGS 60
|||

QY 61 GRAAGKAVSCAIEIVKRRVPGLHQLTKLRFQ 91
|||
Db 61 GRAAGKAVSCAIEIVKRRVPGLHQLTKLRFQ 91
|||


```
RESULT 2
US-09-876-997-459
; Sequence 459, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-997-459

Query Match      100.0%; Score 463; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91

RESULT 3
US-09-978-360A-668
; Sequence 668, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56 US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
```

```
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 668
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52...-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)
; OTHER INFORMATION: unknown
US-09-978-360A-668

Query Match      98.7%; Score 457; DB 11; Length 163;
Best Local Similarity 98.9%; Pred. No. 1.9e-42;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91
Db 61 GRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91

RESULT 4
US-10-425-115-237341
; Sequence 237341, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 237341
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_148039C.1.pep
US-10-425-115-237341

Query Match      65.9%; Score 305; DB 17; Length 72;
Best Local Similarity 86.1%; Pred. No. 4.7e-26;
Matches 62; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFSGS 60

QY 61 GRAAGKAVSCAE 72
Db 61 GRAARNAVSPAE 72

RESULT 5
US-10-424-599-237821
; Sequence 237821, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
```


Db 54 GRAINKTVMIAELIKRRIIVGLHQT 78

RESULT 9

US-10-425-115-254668

; Sequence 254668, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 254668

; LENGTH: 242

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(242)

; OTHER INFORMATION: unsure at all xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_163838C.1.pep

US-10-425-115-254668

Query Match 25.4%; Score 117.5; DB 17; Length 242;

Best Local Similarity 31.0%; Pred. No. 0.00011;

Matches 27; Conservative 14; Mismatches 39; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60

Db 1 MDRYQR-----VEKPRNETPIRENEIRITALGRMRNYIGYMSLLEENGHDEITIKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLK 87

Db 54 GRAINKTVMIAELIKRRIIVGLHQT 80

RESULT 10

US-10-424-599-229595

; Sequence 229595, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 229595

; LENGTH: 245

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_49349C.1.pep

US-10-424-599-229595

Query Match 25.2%; Score 116.5; DB 15; Length 245;

Best Local Similarity 31.0%; Pred. No. 0.00015;

Matches 27; Conservative 16; Mismatches 37; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60

Db 1 MDRYQR-----VEKPKAESPINENEIRVTSQGRTRNYITYATLLQKEGSSEIVLKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTLK 87

Db 54 GRAINKTVMIAELIKRRIIVGLHQT 80

RESULT 11

US-10-437-963-152741

; Sequence 152741, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 152741

; LENGTH: 263

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_52762C.1.pep

US-10-437-963-152741

Query Match 25.2%; Score 116.5; DB 16; Length 263;

Best Local Similarity 29.4%; Pred. No. 0.00016;

Matches 25; Conservative 15; Mismatches 38; Indels 7; Gaps 1;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60

Db 1 MDRYQR-----VEKPRNETPISENEIRITAQGRMRNYISYGMSLLEENGHDEIVIKAM 53

QY 61 GRAAGKAVSCAEIVKRRVPGHLQTL 85

Db 54 GRAINKTVMVVELIKRRIIGGLHQT 78

RESULT 12

US-10-437-963-184232

; Sequence 184232, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 184232

; LENGTH: 261

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(261)

; OTHER INFORMATION: unsure at all xaa locations

Search completed: January 26, 2005, 15:28:36
Job time : 76.8847 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:52:47 ; Search time 23.1356 Seconds
(without alignments)
260.851 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGHLHQLTKLRFLQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	463	100.0	128	4	US-09-621-976-4077 Sequence 4077, Ap
2	395	85.3	78	4	US-09-513-999C-7680 Sequence 7680, Ap
3	151.5	32.7	118	4	US-09-270-767-39818 Sequence 39818, A
4	151.5	32.7	118	4	US-09-270-767-55035 Sequence 55035, A
5	75.5	16.3	261	4	US-09-489-039A-12408 Sequence 12408, A
6	73.5	15.9	271	3	US-08-968-563-33 Sequence 33, Appl
7	73.5	15.9	271	3	US-08-969-683A-33 Sequence 33, Appl
8	73.5	15.9	271	3	US-09-297-928-14 Sequence 14, Appl
9	72.5	15.7	290	4	US-09-252-991A-23012 Sequence 23012, A
10	71	15.3	566	4	US-09-252-991A-18531 Sequence 18531, A
11	70	15.1	284	4	US-09-252-991A-17438 Sequence 17438, A
12	69	14.9	183	4	US-09-800-729-101 Sequence 101, App
13	69	14.9	482	3	US-09-134-001C-4309 Sequence 4309, Ap
14	69	14.9	639	3	US-09-347-801-17 Sequence 17, Appl
15	69	14.9	639	4	US-09-854-731-17 Sequence 17, Appl
16	68.5	14.8	499	4	US-09-270-767-31754 Sequence 31754, A
17	68.5	14.8	499	4	US-09-270-767-46971 Sequence 46971, A
18	68	14.7	482	4	US-09-252-991A-29113 Sequence 29113, A
19	68	14.7	517	4	US-09-252-991A-24083 Sequence 24083, A
20	67.5	14.6	250	3	US-08-968-563-17 Sequence 17, Appl
21	67.5	14.6	250	3	US-08-969-683A-17 Sequence 17, Appl
22	67.5	14.6	250	3	US-09-297-928-13 Sequence 13, Appl
23	67.5	14.6	250	4	US-09-641-652-56 Sequence 56, Appl
24	67.5	14.6	250	4	US-09-538-092-240 Sequence 240, App
25	67	14.5	70	4	US-09-489-847-375 Sequence 375, App
26	67	14.5	402	4	US-09-252-991A-22463 Sequence 22463, A
27	67	14.5	1186	2	US-08-861-464-8 Sequence 8, Appli

28	67	14.5	1186	2	US-08-396-001-8 Sequence 8, Appli
29	67	14.5	1186	3	US-09-323-433A-8 Sequence 8, Appli
30	67	14.5	1186	4	US-09-826-752-8 Sequence 8, Appli
31	67	14.5	2595	3	US-09-036-987A-2 Sequence 2, Appli
32	67	14.5	2595	3	US-09-370-700-2 Sequence 2, Appli
33	67	14.5	2595	4	US-09-603-207-2 Sequence 2, Appli
34	66.5	14.4	496	3	US-09-134-001C-3001 Sequence 3001, Ap
35	66.5	14.4	555	4	US-09-252-991A-28848 Sequence 28848, A
36	66.5	14.4	580	4	US-09-252-991A-27245 Sequence 27245, A
37	65.5	14.1	240	4	US-09-252-991A-31100 Sequence 31100, A
38	65	14.0	877	2	US-08-916-917-2 Sequence 2, Appli
39	65	14.0	877	2	US-08-972-631-2 Sequence 2, Appli
40	65	14.0	877	2	US-08-972-629-2 Sequence 2, Appli
41	65	14.0	877	2	US-08-972-630-2 Sequence 2, Appli
42	65	14.0	877	2	US-08-672-211-2 Sequence 2, Appli
43	65	14.0	877	3	US-09-225-170-2 Sequence 2, Appli
44	64.5	13.9	244	1	US-08-696-827-1 Sequence 1, Appli
45	64.5	13.9	294	4	US-09-252-991A-29716 Sequence 29716, A

ALIGNMENTS

RESULT 1
US-09-621-976-4077
; Sequence 4077, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4077
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52...-1
; NAME/KEY: UNSURE
; LOCATION: 72
; OTHER INFORMATION: Xaa = Asp,Glu,Gly,Val
US-09-621-976-4077

Query Match 100.0%; Score 463; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No.3e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60
|||
Db 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGS 60

QY 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLQ 91
|||
Db 61 GRAAGKAVSCAEIVKRRVPGHLHQLTKLRFLQ 91

RESULT 2
US-09-513-999C-7680
; Sequence 7680, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7680
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7680

Query Match 85.3%; Score 395; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 5.1e-42;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVVFGS 60
Db 1 MEHYRKAGSVELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRV 78
Db 61 GRAAGKAVSCAEIVKRRV 78

RESULT 3

US-09-270-767-39818
; Sequence 39818, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39818
; LENGTH: 118
; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-39818

Query Match 32.7%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 2.8e-11;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVVFGSGRAAGKAVSC 70
Db 1 DLPFEDCMPKSKQDFLWMHVKGTKVSNVIEFAQEAALNKGEHRCVWWSGGGVGKTISC 60
QY 71 AEIVKRRVPGHLQTLKRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 4

US-09-270-767-55035
; Sequence 55035, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55035
; LENGTH: 118
; TYPE: PRT

; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55035

Query Match 32.7%; Score 151.5; DB 4; Length 118;
Best Local Similarity 41.8%; Pred. No. 2.8e-11;
Matches 33; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 11 ELPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVVFGSGRAAGKAVSC 70
Db 1 DLPFEDCMPKSKQDFLWMHVKGTKVSNVIEFAQEAALNKGEHRCVWWSGGGVGKTISC 60
QY 71 AEIVKRRVPGHLQTLKRF 89
Db 61 AEVLKRSHP-LYQVTRMAY 78

RESULT 5

US-09-489-039A-12408
; Sequence 12408, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12408
; LENGTH: 261
; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12408

Query Match 16.3%; Score 75.5; DB 4; Length 261;
Best Local Similarity 27.7%; Pred. No. 0.26;
Matches 23; Conservative 18; Mismatches 31; Indels 11; Gaps 3;

QY 11 ELPAPSPMPQLPPDPTL-----EMRVRDGSKIRNLLGLALGRLEGGSARHVVVFGSGRAAG 65
Db 178 DLPKIAVMPQSPQDVLTLAATLTMKKYATRPLITMSMGK-SGGVSR-----VTGRLLFG 231
QY 66 KAVSCAEIVKRRVPGHLQTLKLR 88
Db 232 SAMFTGTGVQSSAPGQIAIAKLR 254

RESULT 6

US-08-968-563-33
; Sequence 33, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:

; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITE
; APPLICANT: VASANTHA NAGARAJAN
; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

```
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,563
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: NOVEMBER 13, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9982
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: GPP1
; US-08-968-563-33

Query Match 15.9%; Score 73.5; DB 3; Length 271;
Best Local Similarity 28.3%; Pred. No. 0.48;
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;

QY 2 EHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVVF--SG 59
Db 156 EYFITANDVKQKPHPEPYL-----KRNGLGFPINEQDPSKSKVVVFEDAP 202
QY 60 SGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91
Db 203 AGIAGKAAGC-----KIVGIATTFDLDFLK 228

RESULT 7
US-08-969-683A-33
; Sequence 33, Application US/08969683A
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; STREET: 1870 South Winton road
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,683A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20873
; FILING DATE: 13-NOV-1997
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: GPP1
; US-08-969-683A-33

Query Match 15.9%; Score 73.5; DB 3; Length 271;
Best Local Similarity 28.3%; Pred. No. 0.48;
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;

QY 2 EHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVVF--SG 59
Db 156 EYFITANDVKQKPHPEPYL-----KRNGLGFPINEQDPSKSKVVVFEDAP 202
QY 60 SGRAAGKAVSCAEIVKRRVPGHLQTLKRLFLQ 91
Db 203 AGIAGKAAGC-----KIVGIATTFDLDFLK 228

RESULT 8
US-09-297-928-14
; Sequence 14, Application US/09297928
; Patent No. 6358716
; GENERAL INFORMATION:
; APPLICANT: BULTHUIS, BEN A.
; GATENBY, ANTHONY A.
; HAYNIE, SHARON L.
; HSU, AMY K.
; LAREAU, RICHARD D.
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; GLYCEROL BY RECOMBINANT
; ORGANISMS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/297,928
```


; FILING DATE: 11-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/03602
; FILING DATE: NOVEMBER 13, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9981-P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; TELEX: 6717325
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-297-928-14

Query Match 15.9%; Score 73.5; DB 3; Length 271;
Best Local Similarity 28.3%; Pred. No. 0.48;
Matches 26; Conservative 13; Mismatches 32; Indels 21; Gaps 3;

QY 2 EHYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SARHVVVF--SG 59
Db 156 EYFITANDVKQKPHPEPYL-----KGRNGLGFPINEQDPSKSVVWFEDAP 202

QY 60 SGRAAGKAVSCAEIVKRRVPGHLQLTKLRFQ 91
Db 203 AGIAGKAAGC-----KIVGIATTFDLDFLK 228

RESULT 9
US-09-252-991A-23012
; Sequence 23012, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23012
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23012

Query Match 15.7%; Score 72.5; DB 4; Length 290;
Best Local Similarity 34.8%; Pred. No. 0.7;
Matches 23; Conservative 4; Mismatches 30; Indels 9; Gaps 2;

QY 7 AGSVELPAPSPMPQLPPD---TLEMRVDRGSKIRNLLGLALGRLEGG SARHVVVFGSGRA 63
Db 229 ADSPALPVQPEQHRQPDRA GPAEGRACGAGVRPVPVPG-----QPGEARHFSFLGGTG 282

QY 64 AGKAVS 69
Db 283 AGKALS 288

RESULT 10
US-09-252-991A-18531

; Sequence 18531, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18531
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18531

Query Match 15.3%; Score 71; DB 4; Length 566;
Best Local Similarity 28.7%; Pred. No. 2.6;
Matches 25; Conservative 8; Mismatches 18; Indels 36; Gaps 5;

QY 5 RKAGSVELPAP-----SPMPQL---PPDTLEMRVDRGSKIRNLLGLALGRLEGG SARH 54
Db 87 QRRGAVHYPRPRLPAIPGDHPHPRAGSWPPDLRLRLRQ-----124

QY 55 VVFSGS-GRAAGKAVSCAEIVKRRVPG 80
Db 125 --PFGAQTATGRRRRSCQRIV-RQAPG 148

RESULT 11
US-09-252-991A-17438
; Sequence 17438, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17438
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17438

Query Match 15.1%; Score 70; DB 4; Length 284;
Best Local Similarity 29.0%; Pred. No. 1.4;
Matches 18; Conservative 10; Mismatches 32; Indels 2; Gaps 1;

QY 12 LPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGG SAR--HVVFGSGRAAGKAVS 69
Db 26 LPSMDKRTQIPDPSFAARLKQAMAMRNLIKQETLAEAGVSONTIHKLTSKGAQSTRKLI 85

QY 70 CA 71
Db 86 IA 87

RESULT 12
US-09-800-729-101
; Sequence 101, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:

```

; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (146)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-800-729-101

Query Match      14.9%; Score 69; DB 4; Length 183;
Best Local Similarity 32.8%; Pred. No. 1.1;
Matches 20; Conservative 5; Mismatches 16; Indels 20; Gaps 2;

QY 14 APSP---MPQLPDDTLEMRVRDGSKI-----RNLGLALGRLEGGSAR 53
Db 87 APSPVWMPAVEPDVGRGRARPLRLIGEVIIFYCAGSCPRGARTQHGLALARLQGGRX 146
QY 54 H 54
Db 147 H 147

RESULT 13
US-09-134-001C-4309
; Sequence 4309, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4309
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4309

Query Match      14.9%; Score 69; DB 3; Length 482;
Best Local Similarity 32.7%; Pred. No. 3.8;
Matches 17; Conservative 9; Mismatches 22; Indels 4; Gaps 1;

QY 21 LPDDTLEMRVRDGSKIR---NLLGLALGRLEGGSARHVVFGSGGRAAGKAV 68
Db 7 IPPSKGEMNMNDTGDKRLSRTSLIGLVIGSMIGGAFNIQSDMGGHAGGLAI 58

RESULT 14
US-09-347-801-17
; Sequence 17, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
```

```

; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Zea mays
US-09-347-801-17

Query Match      14.9%; Score 69; DB 3; Length 639;
Best Local Similarity 25.6%; Pred. No. 5.4;
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVELPAPSPMPQLPDDTLEM-----RVRDGSKIRNLLGLALGRLEGGS 52
Db 108 KPGGAAANASPSPPRRPQVKRVSSAGLLLGSVLRRKTENLKDKYSLGRRLGQGFGTT 167
QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78
Db 168 HLCV----ERATGKELACKSILKRKL 189

RESULT 15
US-09-854-731-17
; Sequence 17, Application US/09854731
; Patent No. 6794561
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/854,731
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/092,438
; PRIOR FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Zea mays
US-09-854-731-17

Query Match      14.9%; Score 69; DB 4; Length 639;
Best Local Similarity 25.6%; Pred. No. 5.4;
Matches 22; Conservative 15; Mismatches 33; Indels 16; Gaps 2;

QY 5 RKAGSVELPAPSPMPQLPDDTLEM-----RVRDGSKIRNLLGLALGRLEGGS 52
Db 108 KPGGAAANASPSPPRRPQVKRVSSAGLLLGSVLRRKTENLKDKYSLGRRLGQGFGTT 167
QY 53 RHVVFSGSGRAAGKAVSCAEIVKRRV 78
Db 168 HLCV----ERATGKELACKSILKRKL 189

Search completed: January 26, 2005, 15:08:52
Job time : 24.1356 secs
```

This Page Blank (uspto)

QY 13 PAPSPMPQLPPDPTLEMRVD-----GSKIRNLLGLALG 45
Db 31 PAPSPRDSLAPEDLEMFVLDPEDGDLWESIRGQLGPMAG 69

RESULT 3
C69383
conserved hypothetical protein AF1067 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C;Accession: C69383
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: C69383
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-89 <KLE>
A;Cross-references: UNIPROT:O29195; GB:AE001029; GB:AE000782; NID:g2689352; PIDN:AAB9017
C;Superfamily: Archaeal histone-like protein, Alba type

Query Match 12.1%; Score 56; DB 2; Length 89;
Best Local Similarity 37.0%; Pred. No. 63;
Matches 17; Conservative 7; Mismatches 20; Indels 2; Gaps 2;

QY 33 GSK-IRNLLGLALGRLEGG SARHVVFGSGRAAGKAVSCAEIVKRR 77
Db 9 GNKPMVNYVLAVLTQFNSG-ATEVSIKARGRAISRAVDVAEIVKRR 53

RESULT 4
I73095
MHC class II beta chain - slender loris (fragment)
C;Species: Loris tardigradus (slender loris)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I73095; I73099
R;Figueroa, F.; O'hUigin, C.; Tichy, H.; Klein, J.
J. Immunol. 152, 4455-4465, 1994
A;Title: The origin of the primate Mhc-DRB genes and allelic lineages as deduced from th
A;Reference number: I56227; MUID:94209665; PMID:8157963
A;Accession: I73095
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-83 <RES>
A;Cross-references: UNIPROT:Q30547; UNIPROT:Q30543; EMBL:Z27126; NID:g415838; PIDN:CAA81
A;Note: DRB*W2005
A;Accession: I73099
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-83 <RE2>
A;Cross-references: EMBL:Z27161; NID:g415844; PID:g415845
A;Note: DRB*W2006

Query Match 11.8%; Score 54.5; DB 2; Length 83;
Best Local Similarity 26.0%; Pred. No. 84;
Matches 20; Conservative 12; Mismatches 34; Indels 11; Gaps 2;

QY 7 AGSVLPAPSPMPQLPPDPTLEMRVDRGSKIRNLLGLALGRLEGG SARHVVFGSGRAAGK 66
Db 18 AGAV-----PGETPLQPGGERPLRQRG-----GVSGGDGAGAAVRRVLEQPEGRPGGG 66

QY 67 AVSCAEIVKRRVPGHLHQ 83
Db 67 AGRGGHVLQTQLPDLRR 83

RESULT 5
C84802

hypothetical protein At2g38210 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84802
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84802
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <STO>
A;Cross-references: UNIPROT:O80446; GB:AE002093; NID:g3335369; PIDN:AAC27170.1; GSPDB:GN
C;Genetics:
A;Gene: At2g38210
A;Map position: 2

Query Match 11.1%; Score 51.5; DB 2; Length 79;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 16; Conservative 5; Mismatches 18; Indels 1; Gaps 1;

QY 40 LGLALGRLEGG SARHVVFGSGRAAGKAVSCAEIVKRRVP 79
Db 28 VGLA-QMLRGGVIMDVVNAEQARIAEEAGACAVMALERVP 66

RESULT 6
G75509
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75509
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75509
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-90 <WHI>
A;Cross-references: UNIPROT:Q9RWZ8; GB:AE001910; GB:AE000513; NID:g6458198; PIDN:AAF1005
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0517
A;Map position: 1

Query Match 11.1%; Score 51.5; DB 2; Length 90;
Best Local Similarity 35.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 4; Mismatches 14; Indels 13; Gaps 2;

QY 5 RKAGSVEL-----PAPSPMPQLPPDPTLEMRVDRGSKIRNLLGLALGRL 47
Db 27 RKTASTKTAASSNTAPAPQPQHAP-----SGSRAPLILGLALAVL 66

RESULT 7
T28356
ORF MSV195 ALI motif gene family protein - Melanoplus sanguinipes entomopoxvirus
C;Species: Melanoplus sanguinipes entomopoxvirus
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T28356
R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A;Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A;Reference number: Z20484; MUID:99102612; PMID:9847359
A;Accession: T28356
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-87 <AFO>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:08:58 ; Search time 193 Seconds
(without alignments)
271.291 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGLHQLTKLRFQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 239185

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65.5	14.1	89	1	ALB2_ARCFU	Q28323 archaeoglob
2	60	13.0	72	2	Q12835	Q12835 homo sapien
3	58	12.5	71	2	Q9J9W4	Q9J9W4 tadpole ede
4	58	12.5	71	2	Q9J9W8	Q9J9W8 bufo marinu
5	58	12.5	71	2	Q9J9X0	Q9J9X0 bufo marinu
6	58	12.5	71	2	Q9J9X2	Q9J9X2 leptodactyl
7	58	12.5	71	2	Q9J9X6	Q9J9X6 frog virus
8	58	12.5	71	2	Q9J9Y2	Q9J9Y2 bufo marinu
9	58	12.5	71	2	Q9J9Y4	Q9J9Y4 bufo marinu
10	58	12.5	71	2	Q9J9Z6	Q9J9Z6 bufo marinu
11	58	12.5	71	2	Q9J9Z8	Q9J9Z8 rana tempor
12	58	12.5	71	2	Q9JA00	Q9JA00 rana tempor
13	57.5	12.4	74	2	Q9SFC3	Q9SFC3 arabidopsis
14	56	12.1	77	2	Q92MC1	Q92MC1 rhizobium m
15	56	12.1	89	1	ALB1_ARCFU	Q29195 archaeoglob
16	55.5	12.0	64	2	Q8S3X2	Q8S3X2 oryza sativ
17	55.5	12.0	86	2	Q6Z5D5	Q6Z5D5 oryza sativ
18	55.5	12.0	86	2	BAD10112	BAD10112 oryza sat
19	55.5	12.0	91	1	ALBA_METTH	Q27527 methanobact
20	55	11.9	91	2	Q6Z896	Q6Z896 oryza sativ
21	55	11.9	91	2	BAD07703	BAD07703 oryza sat
22	55	11.9	91	2	BAD07903	BAD07903 oryza sat
23	54.5	11.8	83	2	Q30543	Q30543 loris tardi
24	54.5	11.8	83	2	Q30547	Q30547 loris tardi
25	54	11.7	52	1	PH68_HUMAN	Q9uht9 homo sapien
26	54	11.7	52	2	AAH69658	Aah69658 homo sapi
27	54	11.7	52	2	AAH69757	Aah69757 homo sapi
28	54	11.7	52	2	AAH69775	Aah69775 homo sapi
29	53.5	11.6	71	2	Q9J9Y8	Q9J9Y8 bufo bufo u
30	53.5	11.6	71	2	Q9J9Z0	Q9J9Z0 bufo bufo u
31	53.5	11.6	71	2	Q9J9Z2	Q9J9Z2 bufo bufo u

32	53.5	11.6	73	2	Q6Z8S3	Q6Z8S3 oryza sativ
33	53.5	11.6	73	2	BAD10039	BAD10039 oryza sat
34	53	11.4	90	2	Q99PM8	Q99PM8 mus musculu
35	53	11.4	91	2	Q96TZ4	Q96TZ4 neurospora
36	52.5	11.3	59	2	Q6IJF3	Q6IJF3 drosophila
37	52	11.2	88	2	Q6FVB9	Q6FVB9 candida gla
38	52	11.2	91	2	Q86H93	Q86H93 dictyosteli
39	51.5	11.1	79	1	PXL4_ARATH	Q80446 arabidopsis
40	51.5	11.1	90	2	Q9RWZ8	Q9RWZ8 deinococcus
41	51	11.0	71	2	Q9J9W6	Q9J9W6 sheetfish i
42	51	11.0	71	2	Q9J9X4	Q9J9X4 guppyfish i
43	51	11.0	71	2	Q9J9Y0	Q9J9Y0 doctor fish
44	51	11.0	71	2	Q9J9Y6	Q9J9Y6 catfish iri
45	51	11.0	71	2	Q9J9Z4	Q9J9Z4 bohle irido

ALIGNMENTS

RESULT 1						
ALB2_ARCFU						
ID	ALB2_ARCFU	STANDARD;	PRT;	89	AA.	
AC	O28323;					
DT	10-OCT-2003 (Rel. 42, Created)					
DT	10-OCT-2003 (Rel. 42, Last sequence update)					
DT	01-OCT-2004 (Rel. 45, Last annotation update)					
DE	DNA/RNA-binding protein Alba 2 (Afl)					
GN	Name=albA2; OrderedlocusNames=AF1956;					
OS	Archaeoglobus fulgidus.					
OC	Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;					
OC	Archaeoglobaceae; Archaeoglobus.					
OX	NCBI_TaxID=2234;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=VC-16 / DSM 4304 / ATCC 49558;					
RX	MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;					
RA	Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E., Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D., Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C., Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R., Kirkness E.F., Dougherty B.A., Mckenney K., Adams M.D., Loftus B.J., Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A., Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;					
RT	"The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus.";					
RL	Nature 390:364-370(1997).					
RN	[2]					
RP	X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS), AND MUTAGENESIS OF ASN-10; LYS-11; LEU-18 AND PHE-54.					
RX	MEDLINE=22726758; PubMed=12730210; DOI=10.1074/jbc.M303666200;					
RA	Zhao K., Chai X., Marmorstein R.;					
RT	"Structure of a Sir2 substrate, Alba, reveals a mechanism for deacetylation-induced enhancement of DNA binding.";					
RL	J. Biol. Chem. 278:26071-26077(2003).					
CC	-!- FUNCTION: Binds double-stranded DNA tightly but without sequence specificity. It is distributed uniformly and abundantly on the chromosome, suggesting a role in chromatin architecture. However, it does not significantly compact DNA. Binds rRNA and mRNA in vivo. May play a role in maintaining the structural and functional stability of RNA, and, perhaps, ribosomes (By similarity).					
CC	-!- SUBUNIT: Homotetramer.					
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (probable).					
CC	-!- PTM: Acetylated. Deacetylation by the SIR2-homolog deacetylase may regulate its activity (By similarity).					
CC	-!- SIMILARITY: Belongs to the archaeal histone-like Alba family.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its					

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE000968; AAB89299.1; -.
DR PIR; C69494; C69494.
DR PDB; INFH; X-ray; A/B=-.
DR PDB; INFJ; X-ray; A=1-89.
DR TIGR; AFI956; -.
DR HAMAP; MF_01122; -. 1.
DR InterPro; IPR002775; DUF78.
DR Pfam; PF01918; DUF78; 1.
DR ProDom; PD010497; DUF78; 1.
DR TIGRFAMs; TIGR00285; DUF78; 1.
KW 3D-structure; Acetylation; Complete proteome; DNA-binding;
KW RNA-binding.
FT MOD_RES 11 11 N6-acetyllysine (By similarity).
FT MUTAGEN 10 10 N->A: Slight decrease in ability to bind
FT MUTAGEN 11 11 DNA.
FT MUTAGEN 11 11 K->R,Q,M: Less stable tetramers;
FT MUTAGEN 18 18 decreased ability to bind DNA.
FT MUTAGEN 18 18 L->R: Less stable tetramers; decreased
FT MUTAGEN 54 54 ability to bind DNA.
FT MUTAGEN 54 54 F->R: Less stable tetramers; decreased
FT SEQUENCE 89 AA; 9894 MW; 96071A4C48CB8ED7 CRC64;

Query Match 14.1%; Score 65.5; DB 1; Length 89;
Best Local Similarity 41.2%; Pred. No. 55;
Matches 21; Conservative 8; Mismatches 19; Indels 3; Gaps 3;

QY 33 GSK-IRNLLGLALGRLEGGSAHVVFSGSRAAGKAVSCAEIVKRR-VPGL 81
Db 9 GNKPMVNVYVLTTLTQLNEG-ADEVVIKARGRAISRAVDVAEIVRNRMFGV 58

RESULT 2
Q12835 PRELIMINARY; PRT; 72 AA.
AC Q12835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Anion exchanger 3 cardiac isoform (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95008042; PubMed=79233606;
RA Yannoukakos D., Stuart-Tilley A., Fernandez H., Fey P., Duyk G.,
RA Alper S.;
RT "Molecular cloning, expression, and chromosomal localization of two
RT isoforms of the AE3 anion exchanger from human heart."
RL Circ. Res. 75:603-614(1994).
DR EMBL; U05597; AAA50749.1; -.
DR PIR; I38497; I38497.
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 7667 MW; 95DCC64DB1B30718 CRC64;

Query Match 13.0%; Score 60; DB 2; Length 72;
Best Local Similarity 43.6%; Pred. No. 1.6e+02;
Matches 17; Conservative 1; Mismatches 15; Indels 6; Gaps 1;

QY 13 PAPSPMPQPPDPTLEMRVRD-----GSKIRNLLGLALG 45
Db 31 PAPSPRSLAPEDLEMFVLDFFEDGDLWESIRQLGPMAG 69

RESULT 3

Q9J9W4 PRELIMINARY; PRT; 71 AA.
ID Q9J9W4;
AC Q9J9W4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Tadpole edema virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Ranavirus.
OX NCBI_TaxID=100214;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses."
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157680; AAF64591.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQQLPPSPKLRFRTRDNQEV 63

RESULT 4
Q9J9W8 PRELIMINARY; PRT; 71 AA.
ID Q9J9W8;
AC Q9J9W8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuelan iridovirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;
OC Bufo iridoviruses.
OX NCBI_TaxID=105553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses."
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157676; AAF64587.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGQQLPPSPKLRFRTRDNQEV 63

RESULT 5
Q9J9X0 PRELIMINARY; PRT; 71 AA.
ID Q9J9X0;
AC Q9J9X0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

Accession	Source	Length	Score	DB	Mismatches	Indels	Gaps
RA	Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A., Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.; "Comparative studies of piscine and amphibian iridoviruses."	145	301-331	2000			
RL	Arch. Virol. 145:301-331(2000).						
DR	EMBL; AF157668; AAF64579.1; -						
KW	Hypothetical protein.						
FT	NON_TER 1						
FT	NON_TER 71						
SQ	SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;						
Query Match 12.5%; Score 58; DB 2; Length 71;							
Best Local Similarity 39.5%; Pred. No. 2.6e+02;							
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;							
QY	3 HYRKAGSVE---LPAPSPMPQLPPD-TLENMVRDGSKI 36						
Db	26 HHPVSGTQCKLVVAQSPGPQLPPSPKLFRRTRDNQEV 63						
RESULT 8							
Q9J9Y2	PRELIMINARY; PRT; 71 AA.						
AC	Q9J9Y2						
DT	01-OCT-2000 (TrEMBLrel. 15, Created)						
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)						
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)						
DE	Hypothetical protein (Fragment).						
OS	Bufo marinus Venezuelan iridovirus 4.						
OC	Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;						
OC	Bufo iridoviruses.						
OX	NCBI_TaxID=105555;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=20214220; PubMed=10752555;						
RA	Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A., Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.; "Comparative studies of piscine and amphibian iridoviruses."	145	301-331	2000			
RL	Arch. Virol. 145:301-331(2000).						
DR	EMBL; AF157662; AAF64573.1; -						
KW	Hypothetical protein.						
FT	NON_TER 1						
FT	NON_TER 71						
SQ	SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;						
Query Match 12.5%; Score 58; DB 2; Length 71;							
Best Local Similarity 39.5%; Pred. No. 2.6e+02;							
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;							
QY	3 HYRKAGSVE---LPAPSPMPQLPPD-TLENMVRDGSKI 36						
Db	26 HHPVSGTQCKLVVAQSPGPQLPPSPKLFRRTRDNQEV 63						
RESULT 9							
Q9J9Y4	PRELIMINARY; PRT; 71 AA.						
AC	Q9J9Y4						
DT	01-OCT-2000 (TrEMBLrel. 15, Created)						
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)						
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)						
DE	Hypothetical protein (Fragment).						
OS	Bufo marinus Venezuelan iridovirus 6.						
OC	Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;						
OC	Bufo iridoviruses.						
OX	NCBI_TaxID=105556;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=20214220; PubMed=10752555;						
RA	Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A., Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.; "Comparative studies of piscine and amphibian iridoviruses."	145	301-331	2000			
RL	Arch. Virol. 145:301-331(2000).						

DR EMBL; AF157660; AAF64571.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGPQLPPSPKLRFRTRDNQEV 63

RESULT 10
QYJ9Z6 ID Q9J9Z6 PRELIMINARY; PRT; 71 AA.
AC Q9J9Z6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bufo marinus Venezuelan iridovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus;
OC Bufo iridoviruses.
OX NCBI_TaxID=105552;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157648; AAF64559.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGPQLPPSPKLRFRTRDNQEV 63

RESULT 11
QYJ9Z8 ID Q9J9Z8 PRELIMINARY; PRT; 71 AA.
AC Q9J9Z8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rana temporaria United Kingdom iridovirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=105558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157646; AAF64557.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGPQLPPSPKLRFRTRDNQEV 63

RESULT 12
QYJAA0 ID Q9JAA0 PRELIMINARY; PRT; 71 AA.
AC Q9JAA0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rana temporaria United Kingdom iridovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=105557;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20214220; PubMed=10752555;
RA Hyatt A.D., Gould A.R., Zupanovic Z., Cunningham A.A.,
RA Hengstberger S., Whittington R.J., Kattenbelt J., Coupar B.E.H.;
RT "Comparative studies of piscine and amphibian iridoviruses.";
RL Arch. Virol. 145:301-331(2000).
DR EMBL; AF157644; AAF64555.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7845 MW; 88A4EBF82A2EDDC0 CRC64;

Query Match 12.5%; Score 58; DB 2; Length 71;
Best Local Similarity 39.5%; Pred. No. 2.6e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 4; Gaps 2;

QY 3 HYRKAGSVE---LPAPSPMPQLPPD-TLEMRVRDGSKI 36
Db 26 HHPVSGTQECKLVVAQSPGPQLPPSPKLRFRTRDNQEV 63

RESULT 13
QYJF03 ID Q9JF03 PRELIMINARY; PRT; 74 AA.
AC Q9JF03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE F17A17.25 protein (AT3g07910/F17A17_25) (Hypothetical protein).
GN Name=F17A17.25;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
RA Banh J., Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [4]
RN SEQUENCE FROM N.A.
RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RP EMBL; AC013483; AAF21201.1; -
DR EMBL; AY060549; AAL31180.1; -
DR EMBL; AY086394; AAM64461.1; -
DR EMBL; AF412117; AAL06569.1; -
KW Hypothetical protein.
SQ SEQUENCE 74 AA; 7368 MW; 44410B581F4D51ED CRC64;

Query Match 12.4%; Score 57.5; DB 2; Length 74;
Best Local Similarity 32.0%; Pred. No. 3e+02;
Matches 16; Conservative 7; Mismatches 12; Indels 15; Gaps 2;

QY 41 GLALGRLEGGSARHVVFGSGRAAGKAVSCAEIVKRRVPGHLQLTKLRL 90
Db 13 GVAVGGALGG-----AVGAVYGTYEIRVKVPGHLH---KVRFI 47

RESULT 14
Q92MC1
ID Q92MC1 PRELIMINARY; PRT; 77 AA.
AC Q92MC1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE HYPOTHETICAL TRANSMEMBRANE PROTEIN.
GN OrderedLocusNames=R02706; ORFNames=SMC00666;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutery M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Guzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoeelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Guzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutery M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47285.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Complete proteome; Hypothetical protein; Transmembrane.
SQ SEQUENCE 77 AA; 8618 MW; 2CEA7D9103F1EAE CRC64;

Query Match 12.1%; Score 56; DB 2; Length 77;
Best Local Similarity 46.2%; Pred. No. 4.5e+02;
Matches 18; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

QY 15 PSPMPQLPPDTLEMR---VRDGSKIRNLLGLALGRLEGG 50
Db 32 PQPAERNPKQTLPRGQGLRFLSPLRNWLGVAL-MLVGG 69

RESULT 15
ALB1_ARCFU
ID ALB1_ARCFU STANDARD; PRT; 89 AA.
AC O29135;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE DNA/RNA-binding protein Alba 1 (AF2).
GN Name=alba1; OrderedLocusNames=AF1067;
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Uterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Woese C.R., Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- FUNCTION: Binds double-stranded DNA tightly but without sequence
CC specificity. It is distributed uniformly and abundantly on the
CC chromosome, suggesting a role in chromatin architecture. However,
CC it does not significantly compact DNA. Binds rRNA and mRNA in
CC vivo. May play a role in maintaining the structural and functional
CC stability of RNA, and, perhaps, ribosomes (By similarity).
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: Acetylated. Deacetylation by the SIR2-homolog deacetylase may
CC regulate its activity (By similarity).
CC -1- SIMILARITY: Belongs to the archaeal histone-like Alba family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:07:38 ; Search time 156 Seconds
(without alignments)
209.259 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGHLQLTKLRFLQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5.

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 1094936

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	402	86.8	79	2	AAV11950	Aay11950 Human 5'
2	395	85.3	78	3	AAG03599	Aag03599 Human sec
3	119	25.7	74	3	AAG36978	Aag36978 Arabidops
4	67	14.5	70	3	AAV91690	Aay91690 Human sec
5	67	14.5	70	8	ADL71771	Adl71771 Novel hum
6	62.5	13.5	63	8	ADN99364	Adn99364 Novel hum
7	62	13.4	84	8	ADG22762	Adg22762 Cyanophag
8	61.5	13.3	86	3	AAG34486	Aag34486 Arabidops
9	57.5	12.4	71	3	AAG58753	Aag58753 Arabidops
10	57.5	12.4	74	3	AAG15942	Aag15942 Arabidops
11	57.5	12.4	85	4	ABG00727	Abg00727 Novel hum
12	55	11.9	75	5	ABP10404	Abp10404 Human ORF
13	54	11.7	83	4	AAG64347	Aag64347 Cu/Zn SOD
14	53.5	11.6	67	4	ABB43040	Abb43040 Peptide #
15	53.5	11.6	67	4	AAM36870	Aam36870 Peptide #
16	53.5	11.6	67	4	AAM76763	Aam76763 Human bon
17	53.5	11.6	67	4	AAM63945	Aam63945 Human bra
18	53.5	11.6	67	4	ABG58446	Abg58446 Human liv
19	53.5	11.6	88	4	ABG18684	Abg18684 Novel hum
20	53	11.4	81	4	AAM86389	Aam86389 Human imm
21	53	11.4	83	8	ADO20465	Ado20465 Human PRO
22	52.5	11.3	64	4	AAM84673	Aam84673 Human imm
23	52.5	11.3	64	4	AAU57537	Aau57537 Propionib
24	52.5	11.3	64	6	ABM54056	Abm54056 Propionib
25	52.5	11.3	76	3	AAG53244	Aag53244 Arabidops

26	52.5	11.3	77	4	AAG73945	Aag73945 Human col
27	52.5	11.3	77	4	AAU64906	Aau64906 Propionib
28	52.5	11.3	77	6	ABM61425	Abm61425 Propionib
29	52	11.2	71	8	ABO58355	Abo58355 Human gen
30	52	11.2	80	4	AAU53651	Aau53651 Propionib
31	52	11.2	80	6	ABM50170	Abm50170 Propionib
32	52	11.2	81	4	AAM95345	Aam95345 Human rep
33	52	11.2	81	4	ABB96039	Abb96039 Human tes
34	52	11.2	87	4	AAM92262	Aam92262 Human dig
35	51.5	11.1	61	4	AAU53022	Aau53022 Propionib
36	51.5	11.1	61	6	ABM49541	Abm49541 Propionib
37	51.5	11.1	65	3	AAG31725	Aag31725 Arabidops
38	51.5	11.1	79	3	AAG31723	Aag31723 Arabidops
39	51	11.0	72	6	ABP80069	Abp80069 N. gonorr
40	51	11.0	75	4	AAM13691	Aam13691 Peptide #
41	51	11.0	75	4	ABB32622	Abb32622 Peptide #
42	51	11.0	75	4	AAM26091	Aam26091 Peptide #
43	51	11.0	75	4	ABB27471	Abb27471 Human pep
44	51	11.0	75	4	ABB18120	Abb18120 Protein #
45	51	11.0	75	4	AAM65829	Aam65829 Human bon

ALIGNMENTS

RESULT 1
AAV11950
ID AAY11950 standard; protein; 79 AA.
XX
AC AAY11950;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID No: 550.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001232.
XX
PR 01-AUG-1997; 97US-00905144.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153780/13.
XX
N-PSDB; AAX40672.
XX
PT New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
PS Claim 34; Page 648; 675pp; English.
XX
CC AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 79 AA;

Query Match 86.8%; Score 402; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 5.9e-40;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
Db |||||
1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRVP 79
Db |||||
61 GRAAGKAVSCAEIVKRRVP 79

RESULT 2
AAG03599
ID AAG03599 standard; protein; 78 AA.
XX
AC AAG03599;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 7680.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR N-PSDB; AAC03605.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 13; SEQ ID NO 7680; 71pp + Sequence Listing; English.
XX

The present sequence is a polypeptide encoded by one of a large number of
5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
prepared from total human RNAs or polyA+ RNAs derived from 30 different
tissues. EST sequences usually correspond mainly to the 3' untranslated
region (UTR) of the mRNA because they are often obtained from oligo-dT
primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
sequences derived from the 5' ends of mRNAs and even in those cases where
longer cDNA sequences have been obtained, the full 5' UTR is rarely
included. 5' ESTs are derived from mRNAs with intact 5' ends and can
therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
are also used in diagnostic, forensic, gene therapy and chromosome
mapping procedures. They are used to obtain upstream regulatory sequences

CC and to design expression and secretion vectors
XX
SQ Sequence 78 AA;
Query Match 85.3%; Score 395; DB 3; Length 78;
Best Local Similarity 100.0%; Pred. No. 4e-39;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
Db |||||
1 MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGGSARHVVFGS 60
QY 61 GRAAGKAVSCAEIVKRRV 78
Db |||||
61 GRAAGKAVSCAEIVKRRV 78

RESULT 3
AAG36978
ID AAG36978 standard; protein; 74 AA.
XX
AC AAG36978;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45395.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-00301439.
XX

PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

DR WPI; 2004-180094/17.

XX New human secreted nucleic acid, useful for diagnosing and treating

PT neurodegenerative, inflammatory, hyperproliferative, metabolic,

PT reproductive, cardiovascular, respiratory or immunological disorders or

PT diseases.

XX Disclosure; SEQ ID NO 375; 234pp; English.

PS

XX

CC The invention describes an isolated human nucleic acid molecule (I)

CC comprising a polynucleotide having a nucleotide sequence at least 95%

CC identical to: a sequence polynucleotide fragment of SEQ ID NO: X or of

CC the cDNA sequence included in ATCC Deposit No: Z, which is hybridisable

CC to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or

CC epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA

CC sequence included in ATCC Deposit No: Z, which is hybridisable to SEQ ID

CC NO: X, having a biological activity. The nucleic acids and polypeptides,

CC pharmaceutical formulations and kits are useful in diagnosing and

CC treating neurodegenerative diseases states, behavioral disorders,

CC inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's

CC disease, Parkinson's diseases or Huntington's diseases), metabolic

CC disorders, immunological disorders (e.g. arthritis, asthma or AIDS),

CC endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic

CC or muscular disorders (e.g. leukaemia), autoimmune disorders, allergy,

CC cancer, cardiovascular, respiratory or pulmonary disorders, disorders or

CC conditions afflicting connective tissue, skin disorders, CNS disorders,

CC congenital disorders, infectious disorders and gastrointestinal

CC disorders. This is the amino acid sequence of a novel human secreted

CC protein fragment of the invention. Note: This sequence does not appear in

CC the printed specification but is available in electronic format from the

CC US patent office at ftp.seqdata.uspto.gov/seqdata.html?DocID=20040034196.

XX

SQ Sequence 70 AA;

Query Match 14.5%; Score 67; DB 8; Length 70;

Best Local Similarity 41.0%; Pred. No. 4.5;

Matches 16; Conservative 2; Mismatches 9; Indels 12; Gaps 1;

QY 13 PAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGS 51

Db 10 PLPSPAPALAP-----AHSLLGLLLGRMSGSS 36

RESULT 6

ADN99364

ID ADN99364 standard; protein; 63 AA.

XX

AC ADN99364;

XX

DT 29-JUL-2004 (first entry)

XX

DE Novel human protein sequence #180.

XX

KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;

KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;

KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;

KW early aging; hormonal imbalance; ischemic heart disease;

KW ulcerative colitis.

XX

OS Homo sapiens.

XX

PN WO2004038003-A2.

XX

PD 06-MAY-2004.

XX

PF 24-OCT-2003; 2003WO-US033947.

XX

PR 25-OCT-2002; 2002US-0421061P.

PR 25-OCT-2002; 2002US-0421080P.

PR 25-OCT-2002; 2002US-0421552P.

PR 25-OCT-2002; 2002US-0421614P.

PR 30-OCT-2002; 2002US-0422177P.

PR 30-OCT-2002; 2002US-0422178P.

PR 15-NOV-2002; 2002US-0426355P.

PR 15-NOV-2002; 2002US-0426384P.

PR 15-NOV-2002; 2002US-0426394P.

PR 15-NOV-2002; 2002US-0426430P.

PR 15-NOV-2002; 2002US-0426916P.

PR 27-NOV-2002; 2002US-0429224P.

PR 27-NOV-2002; 2002US-0429302P.

PR 27-NOV-2002; 2002US-0429326P.

PR 27-NOV-2002; 2002US-0429651P.

PR 04-DEC-2002; 2002US-0430645P.

PR 04-DEC-2002; 2002US-0430651P.

PR 04-DEC-2002; 2002US-0430657P.

PR 04-DEC-2002; 2002US-0430663P.

PR 04-DEC-2002; 2002US-0430668P.

PR 04-DEC-2002; 2002US-0430684P.

PR 05-DEC-2002; 2002US-0430937P.

PR 05-DEC-2002; 2002US-0430965P.

PR 05-DEC-2002; 2002US-0431458P.

PR 12-DEC-2002; 2002US-0433251P.

PR 13-DEC-2002; 2002US-0433316P.

PR 13-DEC-2002; 2002US-0433318P.

PR 23-DEC-2002; 2002US-0436236P.

PR 03-JAN-2003; 2003US-0437914P.

PR 17-JAN-2003; 2003US-0440820P.

PR 17-JAN-2003; 2003US-0440821P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476621P.

PR 09-JUN-2003; 2003US-0476632P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485217P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 08-JUL-2003; 2003US-0485359P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;

PI Wong JGP, Wu G, Zhang H, Zeng C;

XX

DR WPI; 2004-365511/34.

DR N-PSDB; ADN98580.

XX

PT New nucleic acid molecules, useful in preparing a composition for

PT treating or preventing e.g. inflammatory, CNS, bacterial or viral

PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or

PT ulcerative colitis.

XX

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 13.3%; Score 61.5; DB 3; Length 86;
Best Local Similarity 36.5%; Pred. No. 26;

Matches	19;	Conservative	4;	Mismatches	18;	Indels	11;	Gaps	2;
QY	14	APSPMPQLPPD	TLEMRVRDGS	KIRNLLGLALGR	LEGGSARHVV	FGSGGRAAG	65		
		: :			:				
Db	45	APAVKPTCPTD	TLKLVFGG-----	LGGYGGGEGGG-----	YGGSGGGGG	85			
RESULT 9									
AAG58753									
ID	AAG58753 standard; protein; 71 AA.								
XX									
AC	AAG58753;								
XX									
DT	18-OCT-2000 (first entry)								
XX									
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 75896.								
XX									
KW	Protein identification; signal transduction pathway; metabolic pathway;								
KW	hybridisation assay; genetic mapping; gene expression control; promoter;								
KW	termination sequence.								
XX									
OS	Arabidopsis thaliana.								
XX									
PN	EP1033405-A2.								
XX									
PD	06-SEP-2000.								
XX									
PF	25-FEB-2000; 2000EP-00301439.								
XX									
PR	25-FEB-1999;	99US-0121825P.							
PR	05-MAR-1999;	99US-0123180P.							
PR	09-MAR-1999;	99US-0123548P.							
PR	23-MAR-1999;	99US-0125788P.							
PR	25-MAR-1999;	99US-0126264P.							
PR	29-MAR-1999;	99US-0126785P.							
PR	01-APR-1999;	99US-0127462P.							
PR	06-APR-1999;	99US-0128234P.							
PR	08-APR-1999;	99US-0128714P.							
PR	16-APR-1999;	99US-0129845P.							
PR	19-APR-1999;	99US-0130077P.							
PR	21-APR-1999;	99US-0130449P.							
PR	23-APR-1999;	99US-0130510P.							
PR	23-APR-1999;	99US-0130891P.							
PR	28-APR-1999;	99US-0131449P.							
PR	30-APR-1999;	99US-0132048P.							
PR	30-APR-1999;	99US-0132407P.							
PR	04-MAY-1999;	99US-0132484P.							
PR	05-MAY-1999;	99US-0132485P.							
PR	06-MAY-1999;	99US-0132486P.							
PR	06-MAY-1999;	99US-0132487P.							
PR	07-MAY-1999;	99US-0132863P.							
PR	11-MAY-1999;	99US-0134256P.							
PR	14-MAY-1999;	99US-0134218P.							
PR	14-MAY-1999;	99US-0134219P.							
PR	14-MAY-1999;	99US-0134221P.							
PR	14-MAY-1999;	99US-0134370P.							
PR	18-MAY-1999;	99US-0134768P.							
PR	19-MAY-1999;	99US-0134941P.							
PR	20-MAY-1999;	99US-0135124P.							
PR	21-MAY-1999;	99US-0135353P.							
PR	24-MAY-1999;	99US-0135629P.							
PR	25-MAY-1999;	99US-0136021P.							
PR	27-MAY-1999;	99US-0136392P.							
PR	28-MAY-1999;	99US-0136782P.							
PR	01-JUN-1999;	99US-0137222P.							
PR	03-JUN-1999;	99US-0137528P.							
PR	04-JUN-1999;	99US-0137502P.							
PR	07-JUN-1999;	99US-0137724P.							
PR	08-JUN-1999;	99US-0138094P.							
PR	10-JUN-1999;	99US-0138540P.							
PR	10-JUN-1999;	99US-0138847P.							
PR	14-JUN-1999;	99US-0139119P.							
PR	16-JUN-1999;	99US-0139452P.							
PR	16-JUN-1999;	99US-0139453P.							
PR	17-JUN-1999;	99US-0139492P.							
PR	18-JUN-1999;	99US-0139454P.							
PR	18-JUN-1999;	99US-0139455P.							
PR	18-JUN-1999;	99US-0139456P.							
PR	18-JUN-1999;	99US-0139457P.							
PR	18-JUN-1999;	99US-0139458P.							
PR	18-JUN-1999;	99US-0139459P.							
PR	18-JUN-1999;	99US-0139460P.							
PR	18-JUN-1999;	99US-0139461P.							
PR	18-JUN-1999;	99US-0139462P.							
PR	18-JUN-1999;	99US-0139463P.							
PR	18-JUN-1999;	99US-0139750P.							
PR	18-JUN-1999;	99US-0139763P.							
PR	21-JUN-1999;	99US-0139817P.							
PR	22-JUN-1999;	99US-0139899P.							
PR	23-JUN-1999;	99US-0140353P.							
PR	23-JUN-1999;	99US-0140354P.							
PR	24-JUN-1999;	99US-0140695P.							
PR	28-JUN-1999;	99US-0140823P.							
PR	29-JUN-1999;	99US-0140991P.							
PR	30-JUN-1999;	99US-0141287P.							
PR	01-JUL-1999;	99US-0141842P.							
PR	01-JUL-1999;	99US-0142154P.							
PR	02-JUL-1999;	99US-0142055P.							
PR	06-JUL-1999;	99US-0142390P.							
PR	08-JUL-1999;	99US-0142803P.							
PR	09-JUL-1999;	99US-0142920P.							
PR	12-JUL-1999;	99US-0142977P.							
PR	13-JUL-1999;	99US-0143542P.							
PR	14-JUL-1999;	99US-0143624P.							
PR	15-JUL-1999;	99US-0144005P.							
PR	16-JUL-1999;	99US-0144085P.							
PR	16-JUL-1999;	99US-0144086P.							
PR	19-JUL-1999;	99US-0144325P.							
PR	19-JUL-1999;	99US-0144331P.							
PR	19-JUL-1999;	99US-0144332P.							
PR	19-JUL-1999;	99US-0144333P.							
PR	19-JUL-1999;	99US-0144334P.							
PR	19-JUL-1999;	99US-0144335P.							
PR	20-JUL-1999;	99US-0144352P.							
PR	20-JUL-1999;	99US-0144632P.							
PR	20-JUL-1999;	99US-0144884P.							
PR	21-JUL-1999;	99US-0144814P.							
PR	21-JUL-1999;	99US-0145086P.							
PR	21-JUL-1999;	99US-0145088P.							
PR	22-JUL-1999;	99US-0145085P.							
PR	22-JUL-1999;	99US-0145087P.							
PR	22-JUL-1999;	99US-0145089P.							
PR	22-JUL-1999;	99US-0145192P.							
PR	23-JUL-1999;	99US-0145145P.							
PR	23-JUL-1999;	99US-0145218P.							
PR	23-JUL-1999;	99US-0145224P.							
PR	26-JUL-1999;	99US-0145276P.							
PR	27-JUL-1999;	99US-0145913P.							
PR	27-JUL-1999;	99US-0145918P.							
PR	27-JUL-1999;	99US-0145919P.							
PR	28-JUL-1999;	99US-0145951P.							
PR	02-AUG-1999;	99US-0146386P.							
PR	02-AUG-1999;	99US-0146388P.							
PR	02-AUG-1999;	99US-0146389P.							
PR	03-AUG-1999;	99US-0147038P.							
PR	04-AUG-1999;	99US-0147204P.							
PR	04-AUG-1999;	99US-0147302P.							
PR	05-AUG-1999;	99US-0147192P.							
PR	05-AUG-1999;	99US-0147260P.							
PR	06-AUG-1999;	99US-0147303P.							
PR	06-AUG-1999;	99US-0147416P.							
PR	09-AUG-1999;	99US-0147493P.							
PR	09-AUG-1999;	99US-0147935P.							
PR	10-AUG-1999;	99US-0148171P.							
PR	11-AUG-1999;	99US-0148319P.							
PR	12-AUG-1999;	99US-0148341P.							

[illegible]

PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 37139; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 67 AA;

Query Match 11.6%; Score 53.5; DB 4; Length 67;
Best Local Similarity 30.8%; Pred. No. 1.7e+02;
Matches 16; Conservative 8; Mismatches 11; Indels 17; Gaps 3;

Qy 8 GSVELP-----APSPMPQLPPDPTLEMRVRDGSXIR-----NLLGLALGRLE 48
Db 9 GLIDYPLGWPAVSPKPMPELPP-----MGDNTKSQLAMSANFLGSVLTLLQ 54

Search completed: January 26, 2005, 15:31:21
Job time : 160 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:34:56 ; Search time 143 Seconds
(without alignments)
229.911 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVRRVPGLHQLTKLRFQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 671531

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	305	65.9	72	17	US-10-425-115-237341
2	68	14.7	87	15	US-10-424-599-186121
3	67	14.5	70	15	US-10-351-334-375
4	66.5	14.4	79	17	US-10-425-115-240946
5	61.5	13.3	81	16	US-10-437-963-141539
6	60.5	13.1	75	16	US-10-437-963-135348
7	60	13.0	91	16	US-10-767-701-33700
8	58.5	12.6	86	16	US-10-767-701-35766
9	57.5	12.4	80	17	US-10-425-115-259300
10	57.5	12.4	87	17	US-10-425-115-333500
11	57	12.3	61	17	US-10-425-115-362864
12	57	12.3	86	16	US-10-437-963-147885
13	57	12.3	87	16	US-10-437-963-113150

14	56	12.1	90	15	US-10-424-599-190837	Sequence 190837,
15	56	12.1	91	16	US-10-437-963-107873	Sequence 107873,
16	55	11.9	83	16	US-10-437-963-141106	Sequence 141106,
17	54.5	11.8	76	17	US-10-425-115-221369	Sequence 221369,
18	54.5	11.8	88	17	US-10-425-115-316038	Sequence 316038,
19	54.5	11.8	90	17	US-10-425-115-251264	Sequence 251264,
20	54	11.7	76	17	US-10-425-115-359657	Sequence 359657,
21	54	11.7	82	16	US-10-767-701-39398	Sequence 39398, A
22	54	11.7	86	17	US-10-425-115-197142	Sequence 197142,
23	54	11.7	91	16	US-10-437-963-159456	Sequence 159456,
24	53.5	11.6	59	17	US-10-425-115-294016	Sequence 294016,
25	53.5	11.6	67	9	US-09-864-761-45745	Sequence 45745, A
26	53.5	11.6	70	17	US-10-425-115-351248	Sequence 351248,
27	53.5	11.6	88	16	US-10-437-963-136739	Sequence 136739,
28	53.5	11.6	91	17	US-10-425-115-280710	Sequence 280710,
29	53	11.4	68	17	US-10-425-115-297482	Sequence 297482,
30	53	11.4	76	15	US-10-424-599-211238	Sequence 211238,
31	53	11.4	83	17	US-10-425-115-366522	Sequence 366522,
32	53	11.4	86	16	US-10-437-963-165120	Sequence 165120,
33	53	11.4	90	16	US-10-437-963-161882	Sequence 161882,
34	53	11.4	91	16	US-10-437-963-197812	Sequence 197812,
35	52.5	11.3	55	17	US-10-425-115-260434	Sequence 260434,
36	52.5	11.3	65	16	US-10-437-963-170394	Sequence 170394,
37	52.5	11.3	77	14	US-10-106-698-4719	Sequence 4719, Ap
38	52.5	11.3	80	16	US-10-437-963-173391	Sequence 173391,
39	52.5	11.3	86	15	US-10-335-977-8204	Sequence 8204, Ap
40	52.5	11.3	90	16	US-10-437-963-180208	Sequence 180208,
41	52	11.2	47	16	US-10-437-963-173227	Sequence 173227,
42	52	11.2	48	16	US-10-437-963-120541	Sequence 120541,
43	52	11.2	65	17	US-10-425-115-271602	Sequence 271602,
44	52	11.2	71	14	US-10-029-386-31989	Sequence 31989, A
45	52	11.2	71	17	US-10-425-115-319753	Sequence 319753,

ALIGNMENTS

RESULT 1
US-10-425-115-237341
; Sequence 237341, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 237341
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_148039C.1.pep
US-10-425-115-237341

Query Match	65.9%;	Score 305;	DB 17;	Length 72;
Best Local Similarity	86.1%;	Pred. No. 4.7e-26;		
Matches	62;	Conservative	2;	Mismatches 8; Indels 0; Gaps 0;
QY	1	MEHYRKAGSVELPAPSPMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG SARHVVFGS	60	
Db	1	MEHYRKAGSVELPAPSPPLPQLPPDILEMRVRDGSIRIRNLLGLAQRLEGG SALQVVLGS	60	
QY	61	GRAAGKAVSCAE	72	
Db	61	GRAARNAVSAE	72	

Qy 24 DTLEMRVRDGS-----KIRNLLGLALGRLEGGSAHVVFGSGRAAGKAVSCAEIVKRRV 78
 |||||: ||| | ||||| |||||: |||||
 Db 8 DPLERLLWEGSFEEAARRRRAGLAGGRALGGLAPNV--GRGGRGAGRCYWI--LIKRSV 63

RESULT 6
US-10-437-963-135348
; Sequence 135348, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135348
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37033C.1.pap
US-10-437-963-135348

Query Match	13.18;	Score 60.5;	DB 16;	Length 75;
Best Local Similarity	31.58;	Pred. No. 57;		
Matches 23;	Conservative 12;	Mismatches 27;	Indels 11;	Gaps 5;

QY 73 IVKRRVPGLHQLT 85
::|||:
Db 55 DLQALSPF-HQVT 66

RESULT 7
US-10-767-701-33700
; Sequence 33700, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 33700
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C35716_1.pep
US-10-767-701-33700

Query Match 13.0%; Score 60; DB 16; Length 91;
Best Local Similarity 52.2%; Pred. No. 82;
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Db 35 AGAVEAPAAPRRPPGTRRTR 57

RESULT 8

US-10-767-701-35766

; Sequence 35766, Application US/10767701

; Publication No. US20040172684A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Mol

; TITLE OF INVENTION: Plants and Uses Thereof For Plant I

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 35766

; LENGTH: 86

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C74430_1.pap

US-10-767-701-35766

Query Match 12.6%; Score 58.5; DB 16; Length 86;
Best Local Similarity 32.8%; Pred. No. 1.1e+02;
Matches 19; Conservative 5; Mismatches 25; Indels 9; Gaps 2;

```

RESULT 9
US-10-425-115-259300
; Sequence 259300, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259300
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_1
US-10-425-115-259300

```

	Query Match	12.4%;	Score 57.5;	DB 17;	Length 80;
	Best Local Similarity	23.2%;	Pred. No. 1.3e+02;		
	Matches 19; Conservative 10;	Mismatches 34;	Indels 19;	Gaps 1;	
QY	10 VELPAPSPMPQLPPDITLEMRVRDGSKIRNLLGLALGRLEGG SARHVVVFGSGRAACKAVS	69	:	:	:
Ddb	8 VDLVIPSPSPSTPTTVENYLTRGS-----HRTAQQGGETRVGEARP	48	:	:	:
QY	70 CAEIVKRRVPGLHQLTKLRFLO	91	:	:	:
Ddb	49 CHDLGDSDGARRADRVSKLPYGQ	70	:	:	:

RESULT 10
US-10-425-115-333500
; Sequence 333500, Application US/10425115

```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 333500
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_67265C.1.pep
US-10-425-115-333500
```

```
Query Match      12.4%; Score 57.5; DB 17; Length 87;
Best Local Similarity 34.4%; Pred. No. 1.5e+02;
Matches 22; Conservative 9; Mismatches 18; Indels 15; Gaps 5;
```

```
QY 2 EHYRKAGSVEL---PAPSP---MPQLPP-DTLEMRVDRGSKIRN-----LLGLALGRLEG 49
Db 22 EHKRSSPHCELHTCPAPTAGNMLPHLPWPWSLALPQEEGRGCTSSPVLIGLAVG---G 78

QY 50 GSAR 53
Db 79 GEER 82
```

```
RESULT 11
US-10-425-115-362864
; Sequence 362864, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 362864
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_94106C.1.pep
US-10-425-115-362864
```

```
Query Match      12.3%; Score 57; DB 17; Length 61;
Best Local Similarity 30.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 2; Mismatches 18; Indels 26; Gaps 2;
```

```
QY 14 APSMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGSGRAAGKAVSCAEI 73
Db 6 AGPPCPTLGPST-----ALEVGPCRH-----GPASSRAVPCILGT 39

QY 74 VKRRVP 79
Db 40 VGRAVP 45
```

```
RESULT 12
US-10-437-963-147885
; Sequence 147885, Application US/10437963
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147885
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48370C.1.pep
US-10-437-963-147885
```

```
Query Match      12.3%; Score 57; DB 16; Length 86;
Best Local Similarity 45.7%; Pred. No. 1.6e+02;
Matches 16; Conservative 4; Mismatches 13; Indels 2; Gaps 1;
```

```
QY 45 GRL--EGGSARHVVFGSGRAAGKAVSCAEIVKRR 77
Db 45 GRLGEGGQARNGAASGLGRDAGYRMERGEVDGRQ 79
```

```
RESULT 13
US-10-437-963-113150
; Sequence 113150, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 113150
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(87)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_16966C.1.pep
US-10-437-963-113150
```

```
Query Match      12.3%; Score 57; DB 16; Length 87;
Best Local Similarity 32.2%; Pred. No. 1.7e+02;
Matches 19; Conservative 3; Mismatches 31; Indels 6; Gaps 1;
```

```
QY 3 HYRKAGSVELPAPSPMPQLPPDTLEMRVDRGSKIRNLLGLALGRLEGGSARHVVFGSG 61
Db 21 HCHCAILVSSSPSPPPQTPTPLTSRLSSRPSPLVESLG-----EGRKAEIIQESAGG 73
```

RESULT 14

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:24:36 ; Search time 38 Seconds
(without alignments)
158.814 Million cell updates/sec

Title: US-10-057-813-14_COPY_1_91
Perfect score: 463
Sequence: 1 MEHYRKAGSVELPAPSPMPQ.....EIVKRRVPGHLQLTKLRFIQ 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 301706

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	395	85.3	78	4 US-09-513-999C-7680	Sequence 7680, Ap
2	67	14.5	70	4 US-09-489-847-375	Sequence 375, App
3	52	11.2	69	4 US-09-621-976-6159	Sequence 6159, Ap
4	50.5	10.9	81	4 US-09-371-671B-6	Sequence 6, Appli
5	49.5	10.7	85	3 US-09-430-564-8	Sequence 8, Appli
6	49.5	10.7	85	3 US-09-430-564-14	Sequence 14, Appl
7	49	10.6	86	4 US-09-489-039A-13911	Sequence 13911, A
8	48.5	10.5	87	4 US-09-248-796A-16579	Sequence 16579, A
9	48	10.4	45	3 US-09-018-635-37	Sequence 37, Appl
10	48	10.4	45	4 US-09-912-962-37	Sequence 37, Appl
11	48	10.4	61	4 US-09-540-236-3294	Sequence 3294, Ap
12	48	10.4	71	4 US-09-270-767-58217	Sequence 58217, A
13	48	10.4	80	3 US-09-134-001C-2843	Sequence 2843, Ap
14	47	10.2	59	4 US-09-513-999C-6477	Sequence 6477, Ap
15	47	10.2	60	4 US-09-270-767-37413	Sequence 37413, A
16	47	10.2	60	4 US-09-270-767-52630	Sequence 52630, A
17	47	10.2	77	3 US-08-803-346-62	Sequence 62, Appl
18	47	10.2	77	3 US-08-803-346-63	Sequence 63, Appl
19	47	10.2	87	4 US-09-513-999C-6225	Sequence 6225, Ap
20	47	10.2	90	4 US-09-134-000C-3428	Sequence 3428, Ap
21	46.5	10.0	36	6 5451527-6	Patent No. 5451527
22	46.5	10.0	41	6 5451527-9	Patent No. 5451527
23	46.5	10.0	41	6 5451527-10	Patent No. 5451527
24	46.5	10.0	41	6 5451527-11	Patent No. 5451527
25	46.5	10.0	61	4 US-08-469-260A-507	Sequence 507, App
26	46.5	10.0	61	4 US-08-488-446-507	Sequence 507, App
27	46.5	10.0	61	4 US-08-467-344A-507	Sequence 507, App

28	46.5	10.0	61	4 US-09-583-110-3458	Sequence 3458, Ap
29	46.5	10.0	61	4 US-08-424-550B-507	Sequence 507, App
30	46.5	10.0	64	4 US-09-621-976-6713	Sequence 6713, Ap
31	46.5	10.0	90	4 US-09-621-976-4538	Sequence 4538, Ap
32	46	9.9	53	4 US-09-257-179-88	Sequence 88, Appl
33	46	9.9	76	6 5202239-11	Patent No. 5202239
34	45.5	9.8	72	4 US-09-734-492A-6	Sequence 6, Appli
35	45.5	9.8	76	1 US-08-519-777-22	Sequence 22, Appl
36	45.5	9.8	76	1 US-08-742-035-22	Sequence 22, Appl
37	45.5	9.8	76	2 US-08-777-019-22	Sequence 22, Appl
38	45.5	9.8	76	2 US-08-777-143-22	Sequence 22, Appl
39	45.5	9.8	76	3 US-08-775-414-22	Sequence 22, Appl
40	45.5	9.8	76	3 US-08-931-858E-22	Sequence 22, Appl
41	45.5	9.8	76	3 US-08-981-739-22	Sequence 22, Appl
42	45.5	9.8	76	4 US-09-128-026-22	Sequence 22, Appl
43	45.5	9.8	76	4 US-09-220-616-22	Sequence 22, Appl
44	45.5	9.8	76	4 US-09-220-527-22	Sequence 22, Appl
45	45.5	9.8	76	4 US-09-220-407-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-513-999C-7680
; Sequence 7680, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7680
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7680

Query Match	85.3%	Score 395;	DB 4;	Length 78;
Best Local Similarity	100.0%;	Pred. No. 5.1e-42;		
Matches	78;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDCSKIRNLLGLALGRLEGGSARHVVFSGS	60	
Db	1	MEHYRKAGSVELPAPSPMPQLPPDTLEMRVDCSKIRNLLGLALGRLEGGSARHVVFSGS	60	
QY	61	GRAAGKAVSCAEIVKRRV	78	
Db	61	GRAAGKAVSCAEIVKRRV	78	

RESULT 2
US-09-489-847-375
; Sequence 375, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30

```
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 375
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-375.
```

```
Query Match      14.5%; Score 67; DB 4; Length 70;
Best Local Similarity 41.0%; Pred. No. 0.54;
Matches 16; Conservative 2; Mismatches 9; Indels 12; Gaps 1;

QY 13 PAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGG 51
    ||||| | | | | | | | | | | | | | | | | | |
Db 10 PLPSPAPALAP-----ARSLGLLLGRMSGSS 36
```

```
RESULT 3
US-09-621-976-6159
; Sequence 6159, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6159
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 29
; OTHER INFORMATION: Xaa = Ala,Pro
; NAME/KEY: UNSURE
; LOCATION: 32
; OTHER INFORMATION: Xaa = Cys,Gly,Arg,Ser
US-09-621-976-6159
```

```
Query Match      11.2%; Score 52; DB 4; Length 69;
Best Local Similarity 31.2%; Pred. No. 40;
Matches 15; Conservative 5; Mismatches 22; Indels 6; Gaps 1;

QY 3 HYRKAGSVLPAPSPMPQLPPDPTLEMRVRDGSKIRNLLGLALGRLEGG 50
    | | | | | | | | | | | | | | | | | | | |
Db 26 HHRXPGXHECSAPGPAP-----ARHATRWGGPGQDVHSIYWGREGG 67
```

```
RESULT 4
US-09-371-671B-6
; Sequence 6, Application US/09371671B
; Patent No. 6548743
; GENERAL INFORMATION:
; APPLICANT: Sheen, Jen
; APPLICANT: Chiu, Wan-Ling
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING A
; FILE REFERENCE: 00786/370002
; CURRENT APPLICATION NUMBER: US/09/371,671B
; CURRENT FILING DATE: 1999-08-10
```

```
; PRIOR APPLICATION NUMBER: 60/155,934
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/095,938
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-371-671B-6
```

```
Query Match      10.9%; Score 50.5; DB 4; Length 81;
Best Local Similarity 35.4%; Pred. No. 76;
Matches 17; Conservative 7; Mismatches 21; Indels 3; Gaps 1;

QY 44 LGRLEGGSARHVVFSGSGRAAGKAVSCAEIVKR---RVPGLHQLTKLR 88
    | | | | | | | | | | | | | | | | | | | |
Db 16 LDQCEKDKARVLVHCMGSKSRSPAVVVAYLMKRKGWRLAESHQWVKQR 63
```

```
RESULT 5
US-09-430-564-8
; Sequence 8, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Kay K. Lee-Fruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
; TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS
; FILE REFERENCE: 00246/506002
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-564-8
```

```
Query Match      10.7%; Score 49.5; DB 3; Length 85;
Best Local Similarity 36.6%; Pred. No. 1.1e+02;
Matches 15; Conservative 6; Mismatches 13; Indels 7; Gaps 2;

QY 2 EHYRKAGSV---ELPAPSPMPQLPPDT---LEMVRVDGSK 35
    | | | | | | | | | | | | | | | | | | | |
Db 34 EGFRPSPSLPEPTLPLPLPLPPPPSPSTTAPLPIRPPSGTK 74
```

```
RESULT 6
US-09-430-564-14
; Sequence 14, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Kay K. Lee-Fruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
; TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS
; FILE REFERENCE: 00246/506002
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 85
; TYPE: PRT
```



```

/ STREET: 411 Hackensack Avenue
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/912,962
/ FILING DATE: 25-Jul-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/018,635
/ FILING DATE: 04-FEB-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: David A. Jackson
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ TELEX: 133521
/
/ INFORMATION FOR SEQ ID NO: 37:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 45 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ DESCRIPTION: TRF2
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ SEQUENCE DESCRIPTION: SEQ ID NO: 37:
/
/ US-09-912-962-37

```

Query Match	10.4%;	Score 48;	DB 4;	Length 45;
Best Local Similarity	44.4%;	Pred. No. 73;		
Matches 16;	Conservative	3;	Mismatches 9;	Indels 8;
				Gaps 2;

Qy 49 GGSARHWVFGSGRAAGKAVSCAEIVKRR---VPGL 81
|||: |||||: | : ||
Db 5 GGSS-----DGSGRAAGRAARRSSGRRGRHPEGL 35

RESULT 11
US-09-540-236-3294
; Sequence 3294, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3294
; LENGTH: 61
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3294

Query Match	10.4%;	Score 48;	DB 4;	Length 61;
Best Local Similarity	29.3%;	Pred. No. 1.1e+02;		
Matches 12;	Conservative	9;	Mismatches 20;	Indels 0;
				Gaps 0;

```

Qy      40  LGALGRLEGSARHVVFSGSGRAAGKAVSCAIVVKRRVPG  80
      ||:  ::  ::  ::  ||  ||  ||  ||  ||  ||  ::  ::  ||
Db      16  LGITKEEIQEILIQNAIFSGFTRAMNAAVLDEVYKXKLOG  56

```

```

RESULT 12
US-09-270-767-58217
; Sequence 58217, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and
; FILE REFERENCE: File Reference: 7326-0
; CURRENT APPLICATION NUMBER: US/09/270,
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 58217
; LENGTH: 71
; TYPE: prt
; ORGANISM: Drosophila melanogaster
US-09-270-767-58217

```

Query Match 10.4%; Score 48; DB 4; Length 71;
Best Local Similarity 27.1%; Pred. No. 1.3e+02;
Matches 13; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 25 TLEMRVDGSKIRNLLGLALRLEGGSARHVVFSGSGRAAGKAVSCAE 72
 Dd 3 SIQAALRAGNAATTASGSVSMDORGDALLAAISLGSGBAASGSIIPFAE 50

RESULT 13
US-09-134-001C-2843
; Sequence 2843, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND A
; TITLE OF INVENTION: EPIDERMIDIS FOR D
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,0
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,96
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,77
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2843
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2843

Query Match 10.4%; Score 48; DB 3; Length 80;
Best Local Similarity 44.1%; Pred. No. 1.5e+02;
Matches 15; Conservative 3; Mismatches 14; Indels 2; Gaps 2;

Qy 17 PMPQLPPDTLEMRVRDGSKIRNLLGLALGRLEGG 50
| : | | | : | | | : |
Db 37 PIPQPFPRAVRRHV-DALTIQRLCAGA-GGSLSGG 68

RESULT 14
US-09-513-999C-6477
; Sequence 6477, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,9
; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6477
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6477

Query Match 10.2%; Score 47; DB 4; Length 59;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 5 RKAGSVLPAPSPMPQLPP 23
Db 38 RRPQRRPPPPPPPPPLPP 56

RESULT 15

US-09-270-767-37413
; Sequence 37413, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37413
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-37413

Query Match 10.2%; Score 47; DB 4; Length 60;
Best Local Similarity 31.7%; Pred. No. 1.4e+02;
Matches 13; Conservative 6; Mismatches 14; Indels 8; Gaps 2;

QY 3 HYRKA-----GSVELPAPSPMPQLPPD---TLEMVRDGSK 35
Db 2 HYRTAPSGPAAYELHIPAQFPSCPPGARRSLKYHLPDSSR 42

Search completed: January 26, 2005, 15:36:08
Job time : 39 secs

This Page Blank (uspto)


```

Matches 11; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY      8 TEDSWVPASPDGTGLDPLTVRRHVPVWVLLSRD 40
      | : | : | : | : | : | : | : | : | : |
Db      1197 TPNSCIPGAPCDWIAPEGVERHIPVLFDLNLS 1229

RESULT 3
AF3517
lactoylglutathione lyase (EC 4.4.1.5) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3517
R:Delveschio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AF3517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-160 <KUR>
A;Cross-references: UNIPROT:Q8YDW0; UNIPROT:Q8FXNG; GB:AE008918; PIDN:AAL53305.1; PID:gl
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10064
A;Map position: II
C;Keywords: carbon-sulfur lyase

```

		Query Match	Best Local Similarity	Matches	Score 25.7%; 44.4%;	Pred. No. No. 4.5;	DB 2; Length 160;
					Conservative	Mismatches	Gaps
					2;	Indels	1;
QY	15	ASPDTCGLDPLTVRRHVPVAVWLLSRDP	41	. : :	:		
Dd	113	AQLTGRDDPIEVRLPGTWLAFLHDP	136	- - - PGTWLAF	LHDP		

RESULT 4
G75578
aculeacin A acylase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75578
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75578
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-785 <WHI>
A;Cross-references: UNIPROT:Q9RYQ4; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF1238
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0255
A;Map position: 2

```

Query Match          25.0%; Score 55; DB 2; Length 785;
Best Local Similarity 47.8%; Pred. No. 40;
Matches 11; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

QY      11  SWV--PASPDLTGLDPLTVRRHVP 31
      :|: |||: |||||:
Db      500  ARLANPAAPQTGLDPLVGEVNP 522

```

RESULT 5
T42761
Bassoon protein - rat
N;Alternate names: brain-specific synapse-associated protein

```

C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42761
R;Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex,
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42761
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-3938 <DIE>
A;Cross-references: UNIPROT:O88778; EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G341
A;Experimental source: strain Sprague Dawley; brain
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmitter
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

Query Match      24.8%; Score 54.5; DB 2; Length 3938;
Best Local Similarity 35.0%; Pred. No. 2.8e+02;
Matches 14; Conservative 4; Mismatches 5; Indels 17; Gaps 2;

QY      6 LQTEDSW-----VP-----ASPDTGLDPLTVRR 28
      |||||:|
      |||||:|
Db      2534 LQTEQWEAGRSIGIKRHSMFRLRDACEPESGPDPTVRR 2573

```

RESULT 6
T42730
Bassoon protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42730
R:Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex, J. Cell Biol. 142, 499-509, 1998
A:Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localizes to the presynaptic cytoskeleton
A:Reference number: Z22249; MUID:98345363; PMID:9679147
A:Accession: T42730
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3942 <DIE>
A:Cross-references: UNIPROT:Q88737; EMBL:Y17034; NID:G3413809; PIDN:CAA76598.1; PID:G3413809
A:Experimental source: strain 129 SVJ
C:Genetics:
A:Map position: 9F1
A:Introns: 72/2; 208/3; 505/3; 675/3; 2889/3; 3582/1; 3851/3; 3886/1; 3930/1
A:Note: bassoon
C:Function:
A:Description: may be involved in cytomatrix organization at the site of neurotransmitter release
A:Note: component of the presynaptic cytoskeleton
C:Keywords: coiled coil; zinc finger

Query Match	24.8%;	Score 54.5;	DB 2;	Length 3942;
Best Local Similarity	35.0%;	Pred. No. 2.8e+02;		
Matches 14;	Conservative	4;	Mismatches 5;	Indels 17; Gaps 2;
QY	6	LQTEDSW-----VP-----ASPD TGLDPLTVRR	28	
		:	:	:
		:	:	:
Db	2548	LQTEEQWEAGSGIKKRHSMPRLRDACEPESGDPSTVRR	2587	

```

RESULT 7
D87755
protein T21E12.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: D87755
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites Genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 26, 2005, 14:40:40 ; Search time 43.9186 Seconds
(without alignments)
537.138 Million cell updates/sec
Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOTEDSWVPASPDGTG.....DPLTVRRHVPVAVWLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	220	100.0	163	2	Q8N5L8	Q8n5l8 homo sapien
2	216	98.2	163	2	Q99JH1	Q99jh1 mus musculu
3	216	98.2	163	2	Q9CYX4	Q9cyx4 mus musculu
4	125	56.8	224	2	Q6DGS1	Q6dgs1 brachydanio
5	69	31.4	199	2	Q91WE3	Q91we3 mus musculu
6	68	30.9	199	2	Q9NX88	Q9nrx88 homo sapien
7	68	30.9	199	2	Q9BUL9	Q9bul9 homo sapien
8	67	30.5	286	2	Q96I98	Q96i98 homo sapien
9	67	30.5	553	1	Z324_HUMAN	Q8cld8 yersinia pe
10	66.5	30.2	1192	2	Q8CLD8	Aas62645 yersinia
11	66.5	30.2	1192	2	AAS62645	Q8zde8 yersinia pe
12	66.5	30.2	1275	2	Q8ZD68	Q6nnc7 drosophila
13	64	29.1	206	2	Q6NNC7	Q9v9b6 drosophila
14	64	29.1	206	2	Q9V9B6	Aar96155 drosophil
15	64	29.1	206	2	AAR96155	Q82gr6 streptomyce
16	63.5	28.9	436	2	Q82GR6	Q6m2a1 corynebacte
17	60.5	27.5	490	2	Q6M2A1	Caf20771 corynebac
18	60.5	27.5	490	2	CAF20771	Q8lmv2 oryza sativ
19	59.5	27.0	479	2	Q8LMV2	Q7zyt2 xenopus lae
20	59	26.8	484	2	Q7ZYT2	Q72b09 desulfovibr
21	58.5	26.6	354	2	Q72B09	Aas96307 desulfovi
22	58.5	26.6	354	2	AAS96307	Q8hmi6 carapus ber
23	58	26.4	460	2	Q8HMI6	Q6g0r7 bartonella
24	58	26.4	608	2	Q6G0R7	Q6g4x9 bartonella
25	58	26.4	608	2	Q6G4X9	Q8cjin6 streptomyce
26	58	26.4	4557	2	Q8CJN6	Q89eb5 bradyrhizob
27	57.5	26.1	321	2	Q89EB5	Q89t18 bradyrhizob
28	57.5	26.1	690	2	Q89T18	Q7pth5 anopheles g
29	57	25.9	215	2	Q7PTH5	Q7xt90 oryza sativ
30	57	25.9	262	2	Q7XT90	Q72pw0 leptospira
31	57	25.9	606	2	Q72PW0	

32	57	25.9	606	2	Q8F6D0	Q8f6d0 leptospira
33	57	25.9	606	2	AAS70926	Aas70926 leptospir
34	57	25.9	793	2	Q9NDT3	Q9ndt3 caenorhabdi
35	57	25.9	1576	2	Q8IG10	Q8igl0 caenorhabdi
36	57	25.9	1595	2	Q20888	Q20888 caenorhabdi
37	57	25.9	1610	2	Q8MQ20	Q8mq20 caenorhabdi
38	56.5	25.7	160	2	Q8FXN6	Q8fxn6 bruceella su
39	56.5	25.7	160	2	Q8YDW0	Q8ydw0 bruceella me
40	56.5	25.7	219	2	Q6V273	Q6v273 symbiont ba
41	56.5	25.7	219	2	AAQ91803	Aaq91803 symbiont
42	56.5	25.7	866	2	Q8S2P1	Q8s2p1 oryza sativ
43	56.5	25.7	1460	1	PCX1_MOUSE	Q9qycl mus musculu
44	56.5	25.7	1751	2	Q6ZQ41	Q6zq41 mus musculu
45	56.5	25.7	1751	2	BAC98031	Bac98031 mus muscu

ALIGNMENTS

RESULT 1

ID	Q8N5L8	PRELIMINARY;	PRT;	163 AA.
AC	Q8N5L8;			
DT	01-OCT-2002	(TReMBLrel. 22, Created)		
DT	01-OCT-2002	(TReMBLrel. 22, Last sequence update)		
DT	01-JUN-2003	(TReMBLrel. 24, Last annotation update)		
DE	Chromosome 9	open reading frame 23 protein.		
GN	Name=C9orf23;			
OS	Homo sapiens	(Human)		
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria;	Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RX	MEDLINE=22388257;	PubMed=12477932;		
RA	Strausberg R.L.,	Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D.,	Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F.,	Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F.,	Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L.,	Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M.,	Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J.,	Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S.,	Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A.,	McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S.,	Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K.,	Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J.,	Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M.,	Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W.,	Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C.,	Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I.,	Skalska U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J.,	Marra M.A.;		
RT	"Generation and initial	analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci.	U.S.A. 99:16899-16903 (2002).		
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RA	Strausberg R.;			
RL	Submitted (JUN-2002)	to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC032136;	AAH32136.1; -		
DR	Genew; HGNC:19909;	C9orf23.		
DR	InterPro; IPR011574;	Alba DUF78.		
DR	ProDom; PD010497;	Alba DUF78; 1.		
SQ	SEQUENCE	163 AA; 17631 MW; 442C8727191A0BCE CRC64;		

Query Match 100.0%; Score 220; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 3.7e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPVAVWLLSRDP 41
|||||

Db 85 TKLRFQTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 125

RESULT 2

Q99JH1
ID Q99JH1 PRELIMINARY; PRT; 163 AA.
AC Q99JH1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein (RIKEN cDNA 2810432D09).
GN Name=2810432D09Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3;
RX TISSUE=Skeletal muscle;
RA Ievolella C., Zara I., Lanfranchi G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3;
RX TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3;
RX TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ293897; CAC34588.1; -;
DR EMBL; BC047068; AAH47068.1; -;
DR MGD; MGI:1917211; 2810432D09Rik.
DR InterPro; IPR011574; Alba_DUF78.
DR ProDom; PD010497; Alba_DUF78; 1.
KW Hypothetical protein.
SQ SEQUENCE 163 AA; 17675 MW; 87675201AF87F5B6 CRC64;

Query Match 98.2%; Score 216; DB 2; Length 163;
Best Local Similarity 97.6%; Pred. No. 1.3e-20;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TKLRFQTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
|||||
85 TKLRFQTEDSWVPTSPDGTGLDPLTVRRHVPVAVVLLSRDP 125

RESULT 3

Q9CYX4

ID Q9CYX4 PRELIMINARY; PRT; 163 AA.
AC Q9CYX4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2810432D09 product:hypothetical protein, full
insert sequence.
GN Name=2810432D09Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer."
Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

Db	95	TRLRYSRVREVWQSLPGPTQGTPGEPAASLVLKXNVPLGLAILLSKD 142	
RESULT 8			
Q96I98			
ID	Q96I98	PRELIMINARY; PRT; 286 AA.	
AC	Q96I98;		
DT	01-DEC-2001	(TrEMBLrel. 19, Created)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	ZNF324	protein.	
OS	Homo sapiens	(Human).	
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria;	Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Uterus;		
RX	MEDLINE=22388257;	PubMed=12477932;	
RA	Strausberg R.L.,	Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D.,	Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F.,	Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F.,	Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L.,	Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M.,	Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J.,	Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S.,	Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A.,	McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S.,	Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalón D.K.,	Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahey J.,	Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M.,	Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W.,	Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C.,	Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I.,	Skalska U., Smailus D.E., Schnerch A., Schein J.E.,	
RA	Jones S.J.,	Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Uterus;		
RA	Strausberg R.;		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC007717;	AAH07717.1; --	
DR	GO; GO:0005634;	C:nucleus; IEA.	
DR	GO; GO:0003676;	F:nucleic acid binding; IEA.	
DR	GO; GO:0008270;	F:zinc ion binding; IEA.	
DR	GO; GO:0006355;	P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001909;	KRAB.	
DR	InterPro; IPR001005;	Myb_DNA binding.	
DR	InterPro; IPR007087;	Znf_C2H2.	
DR	Pfam; PF01352;	KRAB; 1.	
DR	SMART; SM00349;	KRAB; 1.	
DR	SMART; SM00355;	Znf_C2H2; 1.	
DR	PROSITE; PS50805;	KRAB; 1.	
DR	PROSITE; PS00037;	MYB 1; UNKNOWN 1.	
DR	PROSITE; PS50157;	ZINC_FINGER_C2H2_2; 1.	
SQ	SEQUENCE	286 AA; 31762 MW; 1FC850BC630DC71B CRC64;	
Query Match	30.5%;	Score 67; DB 2; Length 286;	
Best Local Similarity	46.4%;	Pred. No. 1.8;	
Matches	13; Conservative	2; Mismatches	13; Indels 0; Gaps 0;
QY	9	EDSWVPASPDGTGLDPLTVRRHVPVWVL 36	
Db	57	EPPWVPSGTDLTLSRTTYRRNPGSWSL 84	
RESULT 9			
Z324_HUMAN			
ID	Z324_HUMAN	STANDARD; PRT; 553 AA.	
AC	O75467;		
DT	28-FEB-2003	(Rel. 41, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	05-JUL-2004	(Rel. 44, Last annotation update)	
DE	Zinc finger protein 324	(Zinc finger protein ZF5128).	
GN	Name=ZNF324;		
OS	Homo sapiens	(Human).	
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria;	Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kim B.W.,	Cho H.M., Jun D.Y., Kim Y.H.;	
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: May function as a transcription factor.		
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).		
CC	-!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein family.		
CC	-!- SIMILARITY: Contains 9 C2H2-type zinc fingers.		
CC	-!- SIMILARITY: Contains 1 KRAB domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch)		
CC	or send an email to license@isb-sib.ch .		
CC	-----		
DR	EMBL; AF060503;	AAC33716.1; --	
DR	HSSP; P08047;	ISP2.	
DR	GeneW; HGNC:14096;	ZNF324.	
DR	InterPro; IPR001909;	KRAB.	
DR	InterPro; IPR007087;	Znf_C2H2.	
DR	Pfam; PF01352;	KRAB; 1.	
DR	Pfam; PF00096;	zf-C2H2; 9.	
DR	ProDom; PD000003;	Znf_C2H2; 5.	
DR	SMART; SM00349;	KRAB; 1.	
DR	SMART; SM00355;	Znf_C2H2; 9.	
DR	PROSITE; PS50805;	KRAB; 1.	
DR	PROSITE; PS00028;	ZINC_FINGER_C2H2_1; 9.	
DR	PROSITE; PS50157;	ZINC_FINGER_C2H2_2; 9.	
KW	DNA-binding; Metal-binding; Nuclear protein; Repeat;		
KW	Transcription regulation; Zinc-finger.		
FT	DOMAIN 1	72 KRAB.	
FT	ZN_FING	257 279 C2H2-type 1.	
FT	ZN_FING	285 307 C2H2-type 2.	
FT	ZN_FING	313 335 C2H2-type 3.	
FT	ZN_FING	341 363 C2H2-type 4.	
FT	ZN_FING	369 391 C2H2-type 5.	
FT	ZN_FING	397 419 C2H2-type 6.	
FT	ZN_FING	425 447 C2H2-type 7.	
FT	ZN_FING	453 475 C2H2-type 8.	
FT	ZN_FING	481 503 C2H2-type 9.	
SQ	SEQUENCE	553 AA; 61104 MW; BBCE5C469A13108F CRC64;	
Query Match	30.5%;	Score 67; DB 1; Length 553;	
Best Local Similarity	46.4%;	Pred. No. 3.8;	
Matches	13; Conservative	2; Mismatches	13; Indels 0; Gaps 0;
QY	9	EDSWVPASPDGTGLDPLTVRRHVPVWVL 36	
Db	57	EPPWVPSGTDLTLSRTTYRRNPGSWSL 84	
RESULT 10			
Q8CLD8			
ID	Q8CLD8	PRELIMINARY; PRT; 1192 AA.	
AC	Q8CLD8; Q74SW5;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Hypothetical (Putative membrane protein).		
GN	Name=icmF2; OrderedLocusNames=YP2442, Y1553;		

```
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE013758; AAM85122.1; -.
DR EMBL; AE017136; AAS62645.1; -.
DR InterPro; IPR010623; DUF1215.
DR InterPro; IPR009612; ImcF-related.
DR Pfam; PF06744; DUF1215; 1.
DR Pfam; PF06761; ImcF-related; 1.
KW Hypothetical protein.
SQ SEQUENCE 1192 AA; 135060 MW; C53ECD6EA1D8804F CRC64;

Query Match 30.2%; Score 66.5; DB 2; Length 1192;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 3 LRFLQTEDSWVPA-SPDTGLDPLTVRRHVPVAVW 35
Db 723 LMLLQEEDRWVMAQSLSSSLDPLTLRQDVLVWYL 756

RESULT 11
AAS62645
ID AAS62645 PRELIMINARY; PRT; 1192 AA.
AC AAS62645;
DT 24-MAR-2004 (TrEMBLrel. 27, Created)
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN ICMF2 OR YP2442.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017136; AAS62645.1; -.
SQ SEQUENCE 1192 AA; 135060 MW; C53ECD6EA1D8804F CRC64;

Query Match 30.2%; Score 66.5; DB 2; Length 1192;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 3 LRFLQTEDSWVPA-SPDTGLDPLTVRRHVPVAVW 35
Db 723 LMLLQEEDRWVMAQSLSSSLDPLTLRQDVLVWYL 756
```

```
RESULT 12
Q8ZD68
ID Q8ZD68 PRELIMINARY; PRT; 1275 AA.
AC Q8ZD68;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=YPO2724;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414153; CAC92963.1; -.
DR PIR; AD0332; AD0332.
DR InterPro; IPR010623; DUF1215.
DR InterPro; IPR009612; ImcF-related.
DR Pfam; PF06744; DUF1215; 1.
DR Pfam; PF06761; ImcF-related; 1.
KW Complete proteome.
SQ SEQUENCE 1275 AA; 144249 MW; 791FCBEFDC41FA0B CRC64;

Query Match 30.2%; Score 66.5; DB 2; Length 1275;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 3 LRFLQTEDSWVPA-SPDTGLDPLTVRRHVPVAVW 35
Db 806 LMLLQEEDRWVMAQSLSSSLDPLTLRQDVLVWYL 839

RESULT 13
Q6NNC7
ID Q6NNC7 PRELIMINARY; PRT; 206 AA.
AC Q6NNC7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE RE65722p.
GN Name=CG9422;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Carlson J., Chavez C., Friese E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celnik S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011363; AAR96155.1; -.
SQ SEQUENCE 206 AA; 23162 MW; 8B9FD01C279BC5D9 CRC64;

Query Match 29.1%; Score 64; DB 2; Length 206;
Best Local Similarity 30.0%; Pred. No. 3.1;
Matches 12; Conservative 10; Mismatches 16; Indels 2; Gaps 1;

QY 1 TKLRLQTEDSWVPASPDTGLDPLTVRRHVPVAVWVLLSRD 40
```


Tue Feb 1 05:59:28 2005

us-10-057-813-14_copy_85_125.rup

Page 8

Search completed: January 26, 2005, 15:06:29
Job time : 45.9186 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 14:40:04 ; Search time 40.861 Seconds
(without alignments)
359.949 Million cell updates/sec

Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOQTESWVPASPDG.....DPLTVRRHVPVAVWLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_23Sep04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	220	100.0	163	4	AAG89339	Aag89339 Human sec
2	220	100.0	163	6	ABU00014	Abu00014 Human nov
3	220	100.0	219	8	ADM90988	Adm90988 Human pha
4	194	88.2	163	2	AAY36104	Aay36104 Extended
5	194	88.2	163	8	ADP19412	Adp19412 Human sec
6	68	30.9	211	4	AAM38973	Aam38973 Human pol
7	67	30.5	553	8	ADQ17615	Adq17615 Human sof
8	67	30.5	558	8	ADN99926	Adn99926 Novel hum
9	64	29.1	205	4	ABB64410	Abb64410 Drosophil
10	61.5	28.0	96	5	ABP05282	Abp05282 Human ORF
11	61.5	28.0	586	7	ABO64395	AbO64395 Klebsiell
12	58	26.4	1037	7	ADJ71002	Adj71002 Human hea
13	57	25.9	408	7	ADC36134	Adc36134 Weed cont
14	56.5	25.7	4083	8	ADO13847	Adol3847 Protein e
15	56	25.5	264	4	ABG21326	Abg21326 Novel hum
16	56	25.5	266	2	AAR43576	Aar43576 Bovine ad
17	56	25.5	298	3	AAG51877	Aag51877 Arabidops
18	56	25.5	298	3	AAG16840	Aag16840 Arabidops
19	56	25.5	314	3	AAG51848	Aag51848 Arabidops
20	56	25.5	314	3	AAG16839	Aag16839 Arabidops
21	56	25.5	314	3	AAG51876	Aag51876 Arabidops
22	56	25.5	350	3	AAG16838	Aag16838 Arabidops
23	56	25.5	350	3	AAG51875	Aag51875 Arabidops
24	56	25.5	350	3	AAG51847	Aag51847 Arabidops
25	56	25.5	350	8	ADN72883	Adn72883 Thale cre

26	56	25.5	350	8	ADN72183	Adn72183 Thale cre
27	56	25.5	356	3	AAG51846	Aag51846 Arabidops
28	55.5	25.2	402	6	ABU24856	Abu24856 Protein e
29	55.5	25.2	1379	4	AAM78555	Aam78555 Human pro
30	55.5	25.2	1397	4	AAM79539	Aam79539 Human pro
31	55.5	25.2	1446	5	AAO18739	Aao18739 Human NOV
32	55.5	25.2	1451	6	ABJ39112	Abj39112 Molecule
33	55	25.0	785	7	ADC03509	Adc03509 Deinococc
34	54.5	24.8	98	4	AAU53880	Aau53880 Propionib
35	54.5	24.8	98	6	ABM50399	Abm50399 Propionib
36	54.5	24.8	245	6	ABU44710	Abu44710 Protein e
37	54.5	24.8	1124	7	ADB70297	Adb70297 C. neofor
38	54	24.5	72	8	ABO55760	AbO55760 Human gen
39	54	24.5	346	5	ABG97461	Abg97461 S. kaniha
40	54	24.5	1246	6	ABU32194	Abu32194 Protein e
41	54	24.5	1282	7	ABO63127	AbO63127 Klebsiell
42	53.5	24.3	144	3	AAB23464	Aab23464 Wheat LLS
43	53	24.1	185	4	AAM40741	Aam40741 Human pol
44	53	24.1	215	3	AAG33138	Aag33138 Zea mays
45	53	24.1	239	3	AAG33137	Aag33137 Zea mays

ALIGNMENTS

RESULT 1
AAG89339
ID AAG89339 standard; protein; 163 AA.
XX
AC AAG89339;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 459.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
(GEST) GENSET.
Dumas Milne Edwards J, Bougueleret L, Jobert S;
WPI; 2001-367870/38.
N-PSDB; AAH64942.
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Claim 21; Page 910-911; 921pp; English.
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA

probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention

RESULT 2
ABU00014
ID ABU00014 standard; protein: 163 AA.

Sequences ABG99888-ABG99989 and ABU00010-ABU00433 represent human polypeptides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 163 AA;

RESULT 3
ADM90988
ID ADM90988 standard; protein; 219 AA.

PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Hestir K;
XX
DR WPI; 2004-257410/24.
DR N-PSDB; ADM90779, ADM91197.
XX
PT New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 381; 254pp; English.
XX
CC This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC antianaemic, cardiac, vasotropic, antiasthmatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polypeptide is a human protein of the invention.
XX
SQ Sequence 219 AA;

Query Match 100.0%; Score 220; DB 8; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.2e-22;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVVLLSRDP 41
Db 141 TKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVVLLSRDP 181

RESULT 4
AAAY36104
ID AAY36104 standard; protein; 163 AA.
XX
AC AAY36104;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein sequence, SEQ ID NO. 489.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
XX genetic disease.
OS Homo sapiens.
XX
PN WO9931236-A2.
XX

PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB002122.
XX
PR 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
DR N-PSDB; AAX97788.
XX
PT New isolated human secreted proteins.
XX
PS Claim 9; Page 423-424; 516pp; English.
XX
CC This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases
XX
SQ Sequence 163 AA;

Query Match 88.2%; Score 194; DB 2; Length 163;
Best Local Similarity 87.8%; Pred. No. 3.9e-19;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TKLRFLOTEDSWVPASPDGLDPLTVRRHVPVAVVLLSRDP 41
Db 85 TKLRFLOTEDSWVPXSPDTGLXPLTVRRHVPAXWVLLXDRP 125

RESULT 5
ADP19412
ID ADP19412 standard; protein; 163 AA.
XX
AC ADP19412;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human secreted polypeptide #263.
XX
KW Human; secreted protein; genetic disease.
XX
OS Homo sapiens.
XX
PN US2004110939-A1.
XX
PD 10-JUN-2004.
XX
PF 15-OCT-2001; 2001US-00978360.
XX
PR 17-DEC-1998; 98WO-IB002122.
PR 09-FEB-1999; 99WO-IB000282.
PR 21-JUN-2000; 2000WO-IB000951.
PR 15-SEP-2000; 2000US-00663600.
XX
PA (GEST) GENSET SA.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;

PI Duclert A;
XX WPI; 2004-440404/41.
DR N-PSDB; ADP19007.
XX New isolated polynucleotide encoding secreted polypeptide, useful for
PT gene therapy, or in diagnostic procedures to identify individuals having
PT genetic diseases resulting from abnormal expression of the genes.
XX
PS Claim 2; SEQ ID NO 668; 113pp; English.
XX
CC The invention relates to human cDNA sequences that encode human secreted
CC proteins. The invention also relates to an antibody that specifically
CC binds to a polypeptide of the invention and a method of binding the
CC polypeptide to an antibody. The polynucleotides are useful for expressing
CC the entire secreted proteins which they encode and for distinguishing
CC human tissues and cells from non-human tissues and cells, and for
CC distinguishing between human tissues and cells that do or do not express
CC the polynucleotides comprising the cDNAs. The polynucleotides and
CC polypeptides are useful in forensic procedures or diagnostic procedures
CC to identify individuals with genetic diseases resulting from abnormal
CC expression of the genes corresponding to the cDNAs. The sequences are
CC also useful in gene therapy to control or treat genetic diseases. This
CC sequence represents a human secreted polypeptide of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 163 AA;

Query Match 88.2%; Score 194; DB 8; Length 163;
Best Local Similarity 87.8%; Pred. No. 3.9e-19;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 85 TKLXFLOTEDSWVPXSPDTGLXPLTVRRHVPVAXVWLLXRDP 125

RESULT 6
AAM38973
ID AAM38973 standard; protein; 211 AA.
AC AAM38973;
XX
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2118.
XX
KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US034263.
PF
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-0048725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58129.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Example 4; SEQ ID NO 2118; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with neutropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 211 AA;

Query Match 30.9%; Score 68; DB 4; Length 211;
Best Local Similarity 33.3%; Pred. No. 0.37;
Matches 16; Conservative 10; Mismatches 14; Indels 8; Gaps 2;

QY 1 TKLRFLOTEDSW--VPASPDGTGLDP-----LTVRRHVPVAVVLLSRD 40
Db 107 TLRYRSRVREVWQSLPPGPTQGQTPGEPAASLSVLKNVPLGLALLSKD 154

RESULT 7
ADQ17615
ID ADQ17615 standard; protein; 553 AA.
XX
AC ADQ17615;
XX
XX 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated protein - SEQ ID 432.
DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
XX Homo sapiens.
XX
XX WO2004048938-A2.
PN
XX 10-JUN-2004.
PD
XX 26-NOV-2003; 2003WO-US038193.
PF
XX 26-NOV-2002; 2002US-0429739P.
PR
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX Aziz N, Ginsburg WM, Zlotnik A;
PI WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue

PR 07-JUN-2002; 2002JP-00167239.
XX (SUMO) SUMITOMO CHEM CO LTD.
PA Nakajima H, Mukumoto F, Takaishi M;
PI WPI; 2003-523102/49.
XX N-PSDB; ADC36139.
DR
XX
PT Weed controller metabolism proteins deactivating porphyrinogen oxidase
PT (PPO)-inhibiting herbicides by N-demethylation and their genes, useful
PT e.g. in constructing new breeds of herbicide-resistant plants.
XX
PS Claim 1; SEQ ID NO 1; 812pp; Japanese.
XX
CC The invention relates to a novel DNA encoding a weed controller
CC metabolism protein. A protein of the invention has herbicide activity.
CC The proteins and their encoded genes are useful e.g. in constructing new
CC breeds of herbicide-resistant plants and also in developing various
CC agrochemicals. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 408 AA;
Query Match 25.9%; Score 57; DB 7; Length 408;
Best Local Similarity 32.5%; Pred. No. 30;
Matches 13; Conservative 5; Mismatches 14; Indels 8; Gaps 1;
QY 9 EDSWVPASPDGTGLDPLTVRRHVP-----AVWVLLSRD 40
Db 21 QDRTCPFQPTGYDPLREARPLARVTLTYDGRAIWLVTGRD 60
RESULT 14
AD013847
ID AD013847 standard; protein; 4083 AA.
XX
AC AD013847;
XX
DT 15-JUL-2004 (first entry)
XX
DE Protein encoded by dbv gene cluster ORF17.
XX
KW dbv; gene cluster; Actinomadura sp.; biosynthesis; glycopeptide; A40926;
KW 4-hydroxy-phenylglycine; 3; 5-dihydroxy-phenylglycine;
KW heptapeptide skeleton; chlorination; aromatic residue;
KW beta-hydroxylation; tyrosine; cross-linking; N-acetyl glucuroamine;
KW mannosyl; export; antibiotic.
XX
OS Actinomadura sp. ATCC 39727.
XX
PN EPI413626-A1.
XX
PD 28-APR-2004.
XX
PF 23-OCT-2002; 2002EP-00023597.
XX
PR 23-OCT-2002; 2002EP-00023597.
XX
PA (VICU-) VICURON PHARM INC.
XX
PI Donadio S, Sosio M, Beltrametti F;
XX
DR WPI; 2004-332813/31.
DR N-PSDB; AD013830.
XX
PT Novel isolated polypeptide A40926, comprising polypeptide sequence or
PT open reading frame polypeptide involved in biosynthetic pathway of
PT A409626, useful for synthesizing glycopeptide antibiotic A409626.
XX
PS Claim 28; Page 95-112; 164pp; English.
XX
CC This sequence is encoded by the dbv gene cluster from Nonomuria sp.

CC ATCC39727 (formerly Actinomadura sp. ATCC39727). The 37 proteins encoded
CC by the gene cluster are involved in the biosynthesis of the glycopeptide
CC A40926. dbv ORF 1, 2, 5, and 37 encode polypeptides required for the
CC synthesis of 4-hydroxy-phenylglycine residues of A40926. dbv ORF 30-34
CC and 37 encode polypeptides required for the synthesis of 3, 5-dihydroxy-
CC phenylglycine residues of A40926. dbv ORF 16, 17, 25, 26 and 36 encode
CC polypeptides required for the synthesis of heptapeptide skeleton of
CC A40926. dbv ORF 10 encodes polypeptide required for the chlorination of
CC the aromatic residues of A40926. dbv ORF 28 encodes a polypeptide
CC required for the beta-hydroxylation of the tyrosine residue of amino acid
CC 6 of A40926. dbv ORF 11-14 encode polypeptides required for cross-linking
CC of the aromatic residues of amino acids at positions 2 and 4, 4 and 6, 1
CC and 3, and 5 and 7 of A40926. dbv ORF 9, 23 and 29 encode polypeptides
CC required for addition and formation of N-acetyl glucuroamine residue of
CC A40926. dbv ORF 20 or 27 encode polypeptides required for the attachment
CC of mannosyl residues or N-methylation of A40926. dbv ORF 7, 18, 19, 24
CC and 35 encode polypeptides required for the export of A40926 or its
CC precursors. dbv ORF 3, 4, 6, and 22 encode polypeptides required for
CC regulating the expression of one or more genes of the dbv gene cluster.
CC The dbv gene cluster and the proteins encoded by it are useful for
CC producing glycopeptide antibiotic A40926 or its precursor.
XX
SQ Sequence 4083 AA;
Query Match 25.7%; Score 56.5; DB 8; Length 4083;
Best Local Similarity 40.0%; Pred. No. 5.7e+02;
Matches 16; Conservative 3; Mismatches 12; Indels 9; Gaps 1;
QY 11 SWVPASPDGTGLDPLTVRRH-----VPAVWVLLSRDP 41
Db 2385 AYVTAEEGSLDPPDAVREHLAQLPEFMVPAAVVLLDGVP 2424
RESULT 15
ABG21326
ID ABG21326 standard; protein; 264 AA.
XX
AC ABG21326;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21317.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85513.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 51685; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 264 AA;

Query Match 25.5%; Score 56; DB 4; Length 264;
Best Local Similarity 38.7%; Pred. No. 25;
Matches 12; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

QY 16 SPDTGLDPLTVRRHP-----AVVLLSRD 40
Db 197 APQTLINPMEVREHVPRRHASQCFWVKSGRD 227

Search completed: January 26, 2005, 15:01:07
Job time : 42.861 secs

This Page Blank (uspto)


```
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR APPLICATION NUMBER: 2001-06-08
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-997-459

Query Match      100.0%; Score 220; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 6.2e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
      |||||
Db      85 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 125

RESULT 3
US-09-978-360A-668
; Sequence 668, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 668
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -52...-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36, 47, 54, 65, 70, 103, 107, 108)
; OTHER INFORMATION: unknown
US-09-978-360A-668

Query Match      88.2%; Score 194; DB 11; Length 163;
Best Local Similarity 87.8%; Pred. No. 1.7e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 TKLRFLOTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
      |||||
Db      85 TKLRFLOTEDSWVPXSPDGTGLXPLTVRRHVPVAXWVLLXRD 125

RESULT 4
US-10-723-860-432
; Sequence 432, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 432
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-432

Query Match      30.5%; Score 67; DB 17; Length 553;
Best Local Similarity 46.4%; Pred. No. 3.9;
Matches 13; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY      9 EDSWVPASPDGTGLDPLTVRRHVPVAVVVL 36
      |||||
Db      57 EEPWVPSGTDTTLSRTTVRRRNPGSWSL 84

RESULT 5
US-10-425-115-254668
; Sequence 254668, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 254668
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(242)
; OTHER INFORMATION: unsure at all Xaa locations
```



```
US-10-425-114-54515
; Sequence 54515, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54515
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZNFLMO17184G05_FLI.pep
US-10-425-114-54515

Query Match      29.1%; Score 64; DB 15; Length 315;
Best Local Similarity 43.8%; Pred. No. 5.1;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

Qy 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 126 DTWEPL--EEGLLPLETRRHVSMITVTLKPN 155

RESULT 11
US-10-425-114-58679
; Sequence 58679, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58679
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700577207_FLI.pep
US-10-425-114-58679

Query Match      29.1%; Score 64; DB 15; Length 315;
Best Local Similarity 43.8%; Pred. No. 5.1;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

Qy 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 126 DTWEPL--EEGLLPLETRRHVSMITVTLKPN 155

RESULT 12
US-10-425-114-60913
; Sequence 60913, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60913
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-105-A5_FLI.pep
US-10-425-114-60913

Query Match      29.1%; Score 64; DB 15; Length 325;
Best Local Similarity 43.8%; Pred. No. 5.3;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

Qy 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 136 DTWEPL--EEGLLPLETRRHVSMITVTLKPN 165
```

```
RESULT 13
US-10-156-761-11365
; Sequence 11365, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11365
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11365
```

```
Query Match      28.9%; Score 63.5; DB 14; Length 436;
Best Local Similarity 41.0%; Pred. No. 8.5;
Matches 16; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

Qy 3 LRFLQTEDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
Db 172 LSWIVTEHWRPAAVTVSLAALAV---VPFVWLLLRDHP 207
```

```
RESULT 14
US-10-425-115-308917
; Sequence 308917, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 308917
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(261)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_4479C.1.pep
US-10-425-115-308917

Query Match 28.2%; Score 62; DB 17; Length 261;
Best Local Similarity 43.8%; Pred. No. 7.6;
Matches 14; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
|:| | : || || || | : || |
Db 87 DTWEPL--EEGLPLETRHVSMTITVLSKKP 116

RESULT 15
US-10-767-701-45240
; Sequence 45240, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45240
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7873_1.pep
US-10-767-701-45240

Query Match 28.2%; Score 62; DB 16; Length 275;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 14; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 10 DSWVPASPDGTGLDPLTVRRHVPVAVVLLSRDP 41
|:| | : || || || | : || |
Db 87 DTWEPL--EEGLPLETRHVSMTITVLSKKP 116

Search completed: January 26, 2005, 15:28:37
Job time : 35.1898 secs

This Page Blank (uspto)


```

; LENGTH: 586
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10912

Query Match      28.0%; Score 61.5; DB 4; Length 586;
Best Local Similarity 35.7%; Pred. NO. 2.9;
Matches 15; Conservative 11; Mismatches 13; Indels 3; Gaps 3;

QY      3  LRFLQTEDSWVPAS-PDTGL-DP-LTVRRHVPVWVLLSRDP 41
      || :: || | | : || || :: | : | : | : | : |
Db      141 LRQFVEDNFWLPESQSDIYISDPSTLKEHIDKLPVLTREP 182

RESULT 3
US-09-489-039A-9644
; Sequence 9644, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9644
; LENGTH: 1282
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9644

```

Query Match	24.5%;	Score 54;	DB 4;	Length 1282;
Best Local Similarity	33.3%;	Pred. No. 85;		
Matches	7;	Conservative	8;	Mismatches
			6;	Indels
			0;	Gaps
			0;	
QY	12	WPASPD TGLDPLTVRRH VPA	32	
		::: ::: ::: :::		
Db	690	WLPSAPQLNVNPLTIKOOAEA	710	

RESULT 4
US-09-252-991A-29674
; Sequence 29674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29674
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29674

[illegible]

RESULT 5
US-09-252-991A-31900
; Sequence 31900, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31900
; LENGTH: 400
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31900

Query Match 23.4%; Score 51.5; DB 4; Length 400;
Best Local Similarity 35.6%; Pred.No.46;
Matches 16; Conservative 1; Mismatches 17; Indels 11; Gaps 2;

QY 1 TKLRFLOQTEDS-----WV-----PASPDTCGLDPLTVRRHVPVAV 34
||| ||| : | | | | |
Db 312 TKARHLPTDRSEQRHPWPVRAPGTGPASCSCACSSPRPARRPPTPCAW 356
||| ||| : | | | | |

RESULT 6
US-09-489-039A-12558
; Sequence 12558, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12558
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12558

Query Match	23.4%;	Score 51.5;	DB 4;	Length 452;
Best Local Similarity	37.5%;	Pred. NO. 53;		
Matches 15;	Conservative 4;	Mismatches 14;	Indels 7;	Gaps 2;

RESULT 7
US-09-377-557-10
; Sequence 10, Application US/09377557
; Patent No. 6297055
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Emil M. Jr.
; TITLE OF INVENTION: Amino Acid Decarboxylases
; FILE REFERENCE: BB-1237
; CURRENT APPLICATION NUMBER: US/09/377,557
; CURRENT FILING DATE: 1999-08-19

; EARLIER APPLICATION NUMBER: 60/099,493
; EARLIER FILING DATE: September 8, 1998
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-377-557-10

Query Match 23.4%; Score 51.5; DB 3; Length 497;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 12; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 23 PLTVRRHVPVAVVLLSRD 40
||| ||| ||| ||| ||| |||
Db 471 PLTEERHVDAAWKLL-RD 487

RESULT 8
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 9174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..602
; OTHER INFORMATION: /note= "Represents residues 98 to
; OTHER INFORMATION: 731 of human granulocyte colony-stimulating
; OTHER INFORMATION: factor-receptor."
US-08-419-652-6

Query Match 23.4%; Score 51.5; DB 2; Length 602;
Best Local Similarity 43.5%; Pred. No. 76;

Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;
QY 12 WVPASPDGTGLDPLTVRRHVPVAVW 34
||| | | | | | | | | |
Db 428 WVPEPELGKSELT---HYTIFW 447

RESULT 9
US-09-489-039A-13855
; Sequence 13855, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13855
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13855

Query Match 23.4%; Score 51.5; DB 4; Length 717;
Best Local Similarity 37.5%; Pred. No. 94;
Matches 12; Conservative 3; Mismatches 12; Indels 5; Gaps 2;
QY 9 EDSWVPASPDGTGLDPLTVRRHVPVAVVLLSRD 40
||| | | | | | | | | | | | | | |
Db 276 EDEWIPRP--GTDAALV---AGIAWVLLINED 302

RESULT 10
US-07-923-976-6
; Sequence 6, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte
; TITLE OF INVENTION: Colony-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808

REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1500
TELEFAX: 703-415-1508
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-923-976-6

Query Match 23.4%; Score 51.5; DB 1; Length 771;
Best Local Similarity 43.5%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVAV 34
||| | : | ||| |
Db 547 WVPEPELGKSPLT---HYTIFW 566

RESULT 11
5422248-2
Patent No. 5422248
APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
BENSON, CURTIS M.
TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
STIMULATING FACTOR RECEPTORS
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/6,183
FILING DATE: 15-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 587,329
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 522,952
FILING DATE: 03-APR-1990
APPLICATION NUMBER: 416,306
FILING DATE: 03-OCT-1989
APPLICATION NUMBER: 412,816
FILING DATE: 26-SEP-1989
SEQ ID NO: 2:
LENGTH: 783
5422248-2

Query Match 23.4%; Score 51.5; DB 6; Length 783;
Best Local Similarity 43.5%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVAV 34
||| | : | ||| |
Db 547 WVPEPELGKSPLT---HYTIFW 566

RESULT 12
US-07-923-976-4
Sequence 4, Application US/07923976
Patent No. 5574136
GENERAL INFORMATION:
APPLICANT: Nagata, Shigekazu
APPLICANT: Fukunaga, Rikio
TITLE OF INVENTION: DNA Encoding Granulocyte
Stimulating Factor Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones, Tullar & Cooper, P.C.
STREET: P.O. Box 2266 Eads Station
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,976
FILING DATE: 19920922
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 74539/1990
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 176629/1990
FILING DATE: 03-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00375
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hellwege, James W.
REGISTRATION NUMBER: 28,808
REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1500
TELEFAX: 703-415-1508
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 836 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-923-976-4

Query Match 23.4%; Score 51.5; DB 1; Length 836;
Best Local Similarity 43.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVPVAV 34
||| | : | ||| |
Db 547 WVPEPELGKSPLT---HYTIFW 566

RESULT 13
US-07-923-976-8
Sequence 8, Application US/07923976
Patent No. 5574136
GENERAL INFORMATION:
APPLICANT: Nagata, Shigekazu
APPLICANT: Fukunaga, Rikio
TITLE OF INVENTION: DNA Encoding Granulocyte
Stimulating Factor Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones, Tullar & Cooper, P.C.
STREET: P.O. Box 2266 Eads Station
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,976
FILING DATE: 19920922
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 74539/1990
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 176629/1990
FILING DATE: 03-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00375
FILING DATE: 22-MAR-1991

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 863 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-976-8

Query Match      23.4%; Score 51.5; DB 1; Length 863;
Best Local Similarity 43.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY      12 WVPASPDGTGLDPLTVRRHVPVAVW 34
      ||| | | | | | | | | | | | | | |
Db      547 WVPEPPELGKSPLT---HYTIFW 566

RESULT 14
US-09-134-001C-4956
; Sequence 4956, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4956
; LENGTH: 196
; TYPE: PPT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4956

Query Match      23.2%; Score 51; DB 3; Length 196;
Best Local Similarity 39.0%; Pred. No. 23;
Matches 16; Conservative 4; Mismatches 19; Indels 2; Gaps 2;

QY      1 TKLRFLQTESWVPASPDGTGLDPLTVRRHVPVAVWVLLSRDP 41
      ||| | | | | | | | | | | | | | |
Db      118 TGLRLQLIEQGFSAGQLGKIKTAVTM-VAIIWILLG-DP 156

RESULT 15
US-07-923-976-2
; Sequence 2, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte
; TITLE OF INVENTION: Colony-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-976-2

Query Match      23.0%; Score 50.5; DB 1; Length 837;
Best Local Similarity 43.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY      12 WVPASPDGTGLDPLTVRRHVPVAVW 34
      ||| | | | | | | | | | | | | | |
Db      550 WVPEAPRLGMIPLT---HYTIFW 569

Search completed: January 26, 2005, 15:08:53
Job time : 11.4237 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:44:05 ; Search time 39 Seconds
(without alignments)
101.151 Million cell updates/sec

Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOTEDSWVPASPTG.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 9448

Minimum DB seq length: 0
Maximum DB seq length: 41

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	18.2	36	2 T36994	probable transposa
2	37	16.8	35	2 H84214	hypothetical prote
3	36	16.4	34	2 S46474	Ig heavy chain v r
4	36	16.4	36	1 LBRFGV	light-harvesting p
5	35.5	16.1	38	2 G45095	photosystem I ligh
6	35	15.9	25	2 S03275	photosystem II 22K
7	34	15.5	35	2 A61077	glue protein - rib
8	34	15.5	40	2 F45095	photosystem I ligh
9	33.5	15.2	30	2 E45095	photosystem I ligh
10	33.5	15.2	31	2 I61698	myosin - human (fr
11	33	15.0	18	2 S14661	photosystem I prot
12	33	15.0	24	2 A61505	pollen allergen Ph
13	32.5	14.8	38	2 I70022	kallikrein - mouse
14	32	14.5	20	2 S68341	procathepsin L - g
15	32	14.5	28	2 S20393	trypsin inhibitor
16	32	14.5	35	2 D82224	hypothetical prote
17	31.5	14.3	35	2 S46472	Ig heavy chain v r
18	31	14.1	20	2 JN0252	water-soluble 35K
19	31	14.1	31	2 T29628	hypothetical prote
20	31	14.1	34	2 S46469	Ig heavy chain v r
21	31	14.1	39	2 S26937	Ig heavy chain v r
22	31	14.1	40	2 G69471	hypothetical prote
23	30.5	13.9	38	2 I70018	kallikrein - mouse
24	30	13.6	20	2 PQ0071	T-cell receptor be
25	30	13.6	30	2 S55462	mer5 protein homol
26	30	13.6	35	2 S26887	Ig heavy chain v r
27	30	13.6	35	2 S46473	Ig heavy chain v r
28	30	13.6	35	2 C95037	hypothetical prote
29	30	13.6	35	2 S58708	neutral phosphatas

30	30	13.6	37	2 A86127	hypothetical prote
31	30	13.6	39	2 T28154	histone H3 homolog
32	30	13.6	40	2 G82685	hypothetical prote
33	29.5	13.4	37	2 AH2787	hypothetical prote
34	29	13.2	15	2 A45096	thyrotropin-releas
35	29	13.2	18	2 A54651	insulin-like growt
36	29	13.2	22	2 S67974	apolipoprotein Cb1
37	29	13.2	26	2 S30375	agglutinin - Japan
38	29	13.2	28	2 JX0059	serine proteinase
39	29	13.2	28	2 JX0058	trypsin inhibitor
40	29	13.2	38	2 I65220	dopamine D3 recept
41	29	13.2	41	2 D82691	hypothetical prote
42	28.5	13.0	19	2 S29766	cytochrome c(EDH)
43	28	12.7	17	4 I51887	hypothetical EWSR1
44	28	12.7	19	2 JC2059	homeobox 4 protein
45	28	12.7	19	2 B25213	antifreeze glycopr

ALIGNMENTS

RESULT 1

T36994
probable transposase, truncated section 1 [imported] - Streptomyces coelicolor (fragmen
C;Species: Streptomyces coelicolor
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep-2000
C;Accession: T36994
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21618
A;Accession: T36994
A;Status: preliminary; translated from GB/EMBL/DBSJ
A;Molecule type: DNA
A;Residues: 1-36 <OLI>
A;Cross-references: EMBL:AL109949; PIDN:CAB52908.1; GSPDB:GN00070; SCOEDB:SCJ11.23c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ11.23c
C;Superfamily: Streptomyces coelicolor probable transposase SC6G9.06c

Query Match 18.2%; Score 40; DB 2; Length 36;
Best Local Similarity 40.6%; Pred. No. 1.2e+02;
Matches 13; Conservative 3; Mismatches 8; Indels 8; Gaps 2;

QY 8 TEDSWVPASPTGLDP-----LTVRRHVPVAV 33
Db 2 SEPSW--TAPFTGLSPRCWKLVTLRRQGAV 31

RESULT 2

H84214
hypothetical protein Vng0570h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84214
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: H84214
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-35 <STO>
A;Cross-references: UNIPROT:Q9HRS4; GB:AE004437; NID:g10580164; PIDN:AAG19084.1; GSPDB
C;Genetics:
A;Gene: VNG0570H

Query Match 16.8%; Score 37; DB 2; Length 35;
Best Local Similarity 46.7%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;


```

QY      15  ASPDTGLDPLTVRRH 29
         | : | | | | |
Db      15  AAADQGIDPWTADDH 29

```

RESULT 3
S46474
Ig heavy chain V region (DA-9) - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Jun-2000
C:Accession: S46474
R;Cook, G.P.; Tomlinson, I.M.; Walter, G.; Riethman, H.; Carter, N.P.; Buluwela, L.; Win-
Nature Genet. 7, 162-168, 1994
A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telom-
A:Reference number: S46460; MUID:95004581; PMID:7920635
A:Accession: S46474
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-34 <COO>
A:Cross-references: EMBL:Z29984; NID:G505470; PIDN:CAA82874.1; PID:G1335184
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match	16.4%;	Score 36;	DB 2;	Length 34;
Best Local Similarity	43.8%;	Pred. No. 3.8e+02;		
Matches	7;	Conservative	4;	Mismatches 5; Indels 0; Gaps 0;

QY	11	SWVPASPD	TGLDPLTV	26
			:	: ::
Db	8	SWVRQAPG	KGLEWVSV	23

RESULT 4
LBRFGV
light-harvesting protein B-1015 gamma chain - Rhodopseudomonas viridis
C/Species: Rhodopseudomonas viridis
C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C/Accession: A03455
R/Brunisholz, R.A.; Jay, F.; Suter, F.; Zuber, H.
Biol. Chem. Hoppe-Seyler 366, 87-98, 1985
A/Title: The light-harvesting polypeptides of Rhodopseudomonas viridis. The complete amino acid sequence of the B-1015 gamma chain.
A/Reference number: A90694; MUID: 85225948; PMID:3890891

Query Match 16.4%; Score 36; DB 1; Length 36;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 10; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

Qy	5	F	L	Q	T	E	S	V	P	A	S	P	D	T	G	L	-	D	P	L	T	V	R	R	H	V	P	A	V	V	I	L	S	38	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	1	Y	F	A	A	D	G	S	V	V	P	S	I	S	D	N	L	W	P	E	L	G	I	-	L	G	I	P	T	I	W	I	A	L	34

RESULT 5
G45095
Photosystem I light-harvesting complex chlorophyll a/b protein p22.1/p22 - Chlamydomonas
C;Species: Chlamydomonas reinhardtii
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: G45095
R;Bassi, R.; Soen, S.Y.; Frank, G.; Zuber, H.; Rochaix, J.D.
J. Biol. Chem. 267, 25714-25721, 1992
A;Title: Characterization of chlorophyll a/b proteins of photosystem I from Chlamydomona
A;Reference number: A45095; MUID:93100280; PMID:1464588
A;Accession: G45095
A;Status: preliminary
A;Molecule type: protein

A;Residues: 1-38 <BAS>
A;Cross-references: UNIPROT:Q9S8T9
A;Note: sequence extracted from NCBI backbone (NCBIP:120940)
C;Superfamily: chlorophyll a/b-binding protein
C;Keywords: chloroplast; thylakoid

Query Match 16.1%; Score 35.5; DB 2; Length 38;
Best Local Similarity 33.3%; Pred. No. 5e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 11; Gaps 2;

```

Qy      11 SWVPAS-----PDT-----GLDPLTV 26
         :|:|:|         ||         |
Db      4 NWLPGSDAPAWLPDDLPGNYGFDPLSL 30

```

RESULT 6
S03275
photosystem II 22K protein - spinach (fragment)
C:Species: Spinacia oleracea (spinach)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Aug-1998
C:Accession: S03275
R:Murata, N.; Kajiuura, H.; Fujimura, Y.; Miyao, M.; Murata, T.; Watanabe, A.
Prog. Photosyn. Res. 1, 701-704, 1987
A:Title: Partial amino acid sequences of the proteins of pea and spinach ph
A:Reference number: S03269

Query Match	15.9%;	Score 35;	DB 2;	Length 25;
Best Local Similarity	53.3%;	Pred. No. 3.7e+02;		
Matches 8; Conservative	1;	Mismatches 4;	Indels 2;	Gaps 1;

QY	11	SWVPASPD	TG--LDP	23
		:	:	
Db	8	SWIPAVKGG	NFLDP	22

RESULT 7
A61077
glue protein - ribbed mussel (fragments)
C;Species: Geukensia demissa (ribbed mussel)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: A61077
R;R;Waite, J.H.; Hansen, D.C.; Little, K.T.
J. Comp. Physiol. B 159, 517-525, 1989
A;Title: The glue protein of ribbed mussels (Geukensia demissa): a natural
A;Reference number: A61077; PMID:2481690

Query Match	15.5%	Score 34;	DB 2;	Length 35;
Best Local Similarity	40.0%;	Pred. No. 7.2e+02;		
Matches 6;	Conservative	2;	Mismatches 7;	Indels 0;
				Gaps 0;

Qy	9	EDSWVPASPD	TGLDP	23
		:		
pb	9	QTGYVPGYKOT	GYDP	23

RESULT 8
F45095
photosystem I light-harvesting complex chlorophyll a/b protein p18 - Chlamydomonas reinhardtii
Species: Chlamydomonas reinhardtii

Query Match 14.8%; Score 32.5; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 21 LDPLTVRR-HVPAVVV 35
|:||||| |:|
Db 11 LEDLTVRRIPNPGMWL 26

RESULT 14

S68341
procathepsin L - guinea pig (fragments)
C;Species: Cavia porcellus (guinea pig)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 17-Mar-1999
C;Accession: S68341
R;McDonald, J.K.; Emerick, J.M.C.
Arch. Biochem. Biophys. 323, 409-422, 1995
A;Title: Purification and characterization of procathepsin L, a self-processing zymogen
A;Reference number: S68341; MUID:96063618; PMID:7487106
A;Accession: S68341
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11;12-20 <MCD>

Query Match 14.5%; Score 32; DB 2; Length 20;
Best Local Similarity 53.8%; Pred. No. 7.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 QTEDSWVPASPDT 19
|||:|:|
Db 6 QTLDQAQLPKSVDT 18

RESULT 15

S20393
trypsin inhibitor A - balsam pear
C;Species: Momordica charantia (balsam pear, bitter melon)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S20393
R;Huang, Q.; Liu, S.; Tang, Y.; Zeng, F.; Qian, R.
FEBS Lett. 297, 143-146, 1992
A;Title: Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray crystal structure
A;Reference number: S20393; MUID:92201369; PMID:1551419
A;Accession: S20393
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28 <HUA>
A;Cross-references: UNIPROT:P30709
C;Superfamily: squash trypsin inhibitor ITD I

Query Match 14.5%; Score 32; DB 2; Length 28;
Best Local Similarity 38.5%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 28 RHVPAVWVLLSRD 40
|:|:|:|
Db 1 RSCPRIWNECTRD 13

Search completed: January 26, 2005, 15:53:55
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:36:25 ; Search time 193 Seconds
(without alignments)
122.230 Million cell updates/sec

Title: US-10-057-813-14 COPY 85 125

Perfect score: 220

Sequence: 1 TKLRFLQTEDSWVPASPDG.....DPLTVRRHVPVWVLLSRDP 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 56135

```
Minimum DB seq length: 0
Maximum DB seq length: 41
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :      UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	47.5	21.6	37	2	Q73H28	Q73h28 wolbachia p
2	47.5	21.6	37	2	AAS14438	Aas14438 wolbachia
3	39	17.7	32	2	Q8NG01	Q8ng01 homo sapien
4	39	17.7	32	2	Q88LG0	Q88lg0 pseudomonas
5	39	17.7	32	2	Q8VKN1	Q8vkn1 mycobacteri
6	39	17.7	39	2	Q6LDS8	Q6lds8 oryctolagus
7	39	17.7	39	2	AAA31172	Aaa31172 oryctolag
8	38	17.3	24	2	Q95CH7	Q95ch7 semiliquida
9	38	17.3	24	2	Q95CH9	Q95ch9 altingia gr
10	38	17.3	24	2	Q95CI1	Q95ci1 altingia ta
11	37.5	17.0	26	2	O56554	O56554 hepatitis c
12	37	16.8	21	2	Q85UR5	Q85ur5 strephonema
13	37	16.8	21	2	Q85UR6	Q85ur6 strephonema
14	37	16.8	24	2	Q94P09	Q94p09 altingia ob
15	37	16.8	24	2	Q95CG5	Q95cg5 mytilaria l
16	37	16.8	24	2	Q95CG7	Q95cg7 hamamelis j
17	37	16.8	24	2	Q95CG9	Q95cg9 altingia ch
18	37	16.8	24	2	Q95CH1	Q95ch1 liquidadmar
19	37	16.8	24	2	Q95CH3	Q95ch3 liquidadmar
20	37	16.8	24	2	Q95CH5	Q95ch5 semiliquida
21	37	16.8	24	2	Q95CI3	Q95ci3 altingia ex
22	37	16.8	30	2	Q7T211	Q7t211 oreochromis
23	37	16.8	35	2	Q9HRS4	Q9hrs4 halobacteri
24	37	16.8	37	2	Q9BQ22	Q9bq22 homo sapien
25	37	16.8	41	2	Q8KDX9	Q8kdx9 chlorobium
26	36	16.4	24	2	Q94P10	Q94p10 liquidadmar
27	36	16.4	24	2	Q94P11	Q94p11 liquidadmar
28	36	16.4	28	2	Q9KWQ1	Q9kwq1 sphingomona
29	36	16.4	28	2	Q9KWQ4	Q9kwq4 blastomonas
30	36	16.4	33	2	Q863F7	Q863f7 galago cras
31	36	16.4	34	2	Q9HF30	Q9hf30 cryptococcu

RESULT 1

Q73H28	PRELIMINARY;	PRT;	37 AA.
ID	Q73H28		
AC	Q73H28;		
DT	05-JUL-2004	(TREMBLrel. 27, Created)	
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)	
DE	Hypothetical protein.		
GN	OrderedLocusNames=WD0747;		
OS	Wolbachia pipientis wMel.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Rickettsiaceae; Wolbachieae; Wolbachia.		
OC	Rickettsiaceae; Wolbachieae; Wolbachia.		
OX	NCBI_TaxID=66077;		
FN	[1]		

Query Match	21.6%	Score 47.5;	DB 2;	Length 37;
Best Local Similarity	38.7%	Pred. No. 70;		
Matches 12;	Conservative	6;	Mismatches 10;	Indels 3;
			Gaps 2;	

Qy	6	LQTEDSWVPASPD TGLDPLTVRRHVP AVMVL	36
		: : : : : : : :	
Db	9	LTSKISWIPVS- STGMTPTYGHRN-- DALWL	36

RESULT 2

AAS14438
ID AAS14438 PRELIMINARY; PRT; 37 AA.
AC AAS14438;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 13-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN WD0747.
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP SEQUENCE FROM N.A.

AC Q6LDS8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ca2+ ATPase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast twitch muscle;
RX MEDLINE=88169602; PubMed=2965149;
RA Korczak B., Zarain-Herzberg A., Brandl C.J., Ingles C.J., Green N.M.,
RA MacLennan D.H.;
RT "Structure of the rabbit fast-twitch skeletal muscle Ca2+-ATPase
RT gene.";
RL J. Biol. Chem. 263:4813-4819(1988).
DR EMBL; M20531; AAA31172.1; -.
FT NON TER 39
SQ SEQUENCE 39 AA; 4366 MW; F2D5C948C7A3A306 CRC64;

Query Match 17.7%; Score 39; DB 2; Length 39;
Best Local Similarity 40.0%; Pred. No. 1e+03;
Matches 10; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 8 TED--SWVPASPTGLDPLTVRRHV 30
||: :: | || | | ||:
Db 9 TEECLAYFGVSETTGLTPDQVKRHL 33

RESULT 7
AAA31172
ID AAA31172 PRELIMINARY; PRT; 39 AA.
AC AAA31172;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Ca2+ ATPase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=fast twitch muscle;
RX MEDLINE=88169602; PubMed=2965149;
RA Korczak B., Zarain-Herzberg A., Brandl C.J., Ingles C.J., Green N.M.,
RA MacLennan D.H.;
RT "Structure of the rabbit fast-twitch skeletal muscle Ca2+-ATPase
RT gene.";
RL J. Biol. Chem. 263:4813-4819(1988).
DR EMBL; M20531; AAA31172.1; -.
FT NON TER 39
SQ SEQUENCE 39 AA; 4366 MW; F2D5C948C7A3A306 CRC64;

Query Match 17.7%; Score 39; DB 2; Length 39;
Best Local Similarity 40.0%; Pred. No. 1e+03;
Matches 10; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 8 TED--SWVPASPTGLDPLTVRRHV 30
||: :: | || | | ||:
Db 9 TEECLAYFGVSETTGLTPDQVKRHL 33

RESULT 8
Q95CH7
ID Q95CH7 PRELIMINARY; PRT; 24 AA.
AC Q95CH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PsaA (Fragment).

GN Name=psaA;
OS Semiliquidambar cathayensis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Hamamelidaceae; Semiliquidambar.
OX NCBI_TaxID=99446;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304536; AAL13187.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER 24
SQ SEQUENCE 24 AA; 2812 MW; 84E1CC186B6377E3 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 7.9e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVWVLLSRDP 41
: : | | | | : : | | |
Db 1 MIIRSXEPEVKILVDRDP 18

RESULT 9
Q95CH9
ID Q95CH9 PRELIMINARY; PRT; 24 AA.
AC Q95CH9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PsaA (Fragment).
GN Name=psaA;
OS Altingia gracilipes var. serrulata.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Altingiaceae; Altingia.
OX NCBI_TaxID=166113;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304535; AAL13185.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER 24
SQ SEQUENCE 24 AA; 2812 MW; 84E1CC186B6377E3 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 7.9e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 24 LTVRRHVPVWVLLSRDP 41
: : | | | | : : | | |
Db 1 MIIRSXEPEVKILVDRDP 18

RESULT 10
Q95CI1
ID Q95CI1 PRELIMINARY; PRT; 24 AA.
AC Q95CI1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PsaA (Fragment).
GN Name=psaA;
OS Altingia takhtajanii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;


```
DR EMBL; AF304541; AAL13197.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER 24 24
SQ SEQUENCE 24 AA; 2799 MW; 84E1C2C3EB6377E3 CRC64;

Query Match 16.8%; Score 37; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 1.1e+03;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 24 LTVRRHVPVAVVLLSRDP 41
Db 1 MIIRSPEPEVKILVDRDP 18

RESULT 15
Q95CG5
ID Q95CG5 PRELIMINARY; PRT; 24 AA.
AC Q95CG5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE PsaA (Fragment).
GN Name=psaA;
OS Mytilaria laosensis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Hamamelidaceae; Mytilaria.
OX NCBI_TaxID=54454;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi S., Huang Y., Zhang Q., Boufford D.E., Parks C.R., Wen J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF304544; AAL13203.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER 24 24
SQ SEQUENCE 24 AA; 2799 MW; 84E1C2C3EB6377E3 CRC64;

Query Match 16.8%; Score 37; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 1.1e+03;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 24 LTVRRHVPVAVVLLSRDP 41
Db 1 MIIRSPEPEVKILVDRDP 18
```

Search completed: January 26, 2005, 15:53:20
Job time : 204 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:35:31 ; Search time 157 Seconds
(without alignments)
93.681 Million cell updates/sec

Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOTEDSWVPASPTG.....DPLTVRRHVPVAVVLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 856990

Minimum DB seq length: 0
Maximum DB seq length: 41

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	22.3	37	3	AAB52138
2	49	22.3	37	6	ABP99813 Human sec
3	49	22.3	37	6	ABR01305 Human gen
4	49	22.3	37	6	ADA98381 Human sec
5	48.5	22.0	40	8	ABO55642 Human gen
6	45	20.5	34	3	AAG57081 Arabidops
7	44	20.0	41	4	AAM96628 Human rep
8	44	20.0	41	4	AAU22904 Human pro
9	42.5	19.3	25	4	AAG66872 Human Erb
10	42.5	19.3	32	2	AAW59346 Human Erb
11	42.5	19.3	32	2	AAAY14575 Human erb
12	42.5	19.3	32	4	AAB62075 ErbB2 7C2
13	42.5	19.3	32	4	AAB48764 Human Erb
14	42	19.1	27	7	ADB84514 MSRV-1B p
15	41.5	18.9	15	3	AAAY98932 HLA class
16	41.5	18.9	15	4	AAG89066 HER2/neu
17	41.5	18.9	15	4	AAG88752 HER2/NEU
18	40.5	18.4	32	4	AAM99880 Human exc
19	40.5	18.4	32	4	AAM43679 Human bla
20	40.5	18.4	32	8	ADF71638 Human bla
21	40	18.2	38	4	AAM21828 Peptide #
22	40	18.2	38	4	ABB44199 Peptide #
23	40	18.2	38	4	AAM38152 Peptide #
24	40	18.2	38	4	ABB27077 Protein #
25	40	18.2	38	4	AAM77933 Human bon

26	40	18.2	38	4	AAM65230	Aam65230 Human bra
27	40	18.2	38	4	ABG59572	Abg59572 Human liv
28	40	18.2	38	5	ABG46955	Abg46955 Human pep
29	40	18.2	40	4	AAM94835	Aam94835 Human rep
30	39	17.7	37	3	AAB44942	Aab44942 Human sec
31	39	17.7	40	8	ADK01949	Adk01949 Hepatitis
32	39	17.7	41	5	ABB53187	Abb53187 Human ORF
33	38.5	17.5	28	2	AAR13378	Aar13378 Adult T-c
34	38.5	17.5	33	2	AAR69694	Aar69694 Hepatitis
35	38.5	17.5	33	2	AAR89566	Aar89566 Hepatitis
36	38	17.3	15	3	AAAY98841	Aay98841 HLA class
37	38	17.3	15	4	AAG89011	Aag89011 Her2/neu
38	38	17.3	15	4	AAG88332	Aag88332 HER2/NEU
39	38	17.3	19	4	AAB68622	Aab68622 HER-2 B c
40	38	17.3	32	8	ADH39946	Adh39946 Inhibitor
41	38	17.3	36	4	AAM14155	Aam14155 Peptide #
42	38	17.3	36	4	AAM92083	Aam92083 Human dig
43	38	17.3	36	4	ABB33102	Abb33102 Peptide #
44	38	17.3	36	4	AAM26563	Aam26563 Peptide #
45	38	17.3	36	4	ABB27928	Abb27928 Human pep

ALIGNMENTS

RESULT 1

AAB52138
ID AAB52138 standard; peptide; 37 AA.
XX
AC AAB52138;
XX
DT 22-FEB-2001 (first entry)
XX
DE Human secreted protein encoded by cDNA #36.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnarary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200061624-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US008980.
XX
PR 09-APR-1999; 99US-0128700P.
PR 20-JAN-2000; 2000US-0176930P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-656324/63.
DR N-PSDB; AAC96935.
XX
PT New nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
PS Claim 11; Page 439; 478pp; English.
XX

CC The invention relates to the isolation of genes AAC96900-C96947 encoding
CC the human secreted proteins AAB52104-B52150. This sequence represents a
CC fragment of the protein encoded by the gene given in the descriptor line.
CC The sequence is used as a query sequence for doing BLASTX searches to
CC determine homologous sequence to the protein. The genes and proteins are
CC useful for preventing, ameliorating or treating medical conditions, e.g.
CC by protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,

CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
CC such as myocardial ischaemias; (d) wound healing; (e) neurological
CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
CC such as viral, bacterial, fungal and parasitic infections
XX
SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 3; Length 37;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 26 VRRHVPVWVL 36
:|||||:||||
Db 3 LRRHFPALWVL 13

RESULT 2
ABP99813
ID ABP99813 standard; protein; 37 AA.

XX ABP99813;

DT 26-MAR-2003 (first entry)

DE Human secreted protein SEQ ID NO 757.

XX Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antisickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy.

XX Homo sapiens.

PN WO200277186-A2.

XX 03-OCT-2002.

PF 26-MAR-2002; 2002WO-US009188.

XX 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2003-040583/03.

DR N-PSDB; ABZ67234.

XX New human secreted proteins encoded by genes contained in cDNA clones

PT (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,

PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or

PT West Nile fever.

XX Claim 1; Page 1490; 2423pp; English.

XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the

CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,

CC treating or ameliorating medical conditions e.g. by protein or gene

CC therapy. The genes are isolated from a range of human tissues disclosed

CC in the specification. The nucleic acids, proteins, antibodies and

CC (ant)agonists are useful in the diagnosis, treatment and prevention of:

CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX

SQ Sequence 37 AA;

Query Match 22.3%; Score 49; DB 6; Length 37;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 26 VRRHVPVWVL 36
:|||||:||||
Db 3 LRRHFPALWVL 13

RESULT 3
ABR01305

ID ABR01305 standard; peptide; 37 AA.

XX ABR01305;

DT 12-MAY-2003 (first entry)

DE Human gene 359-encoded secreted protein HTXNZ07, SEQ ID NO:786.

XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnery; chromosome 3p21.31.

XX Homo sapiens.

PN WO200277013-A2.

XX 03-OCT-2002.

PF 26-MAR-2002; 2002WO-US009370.

XX 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2003-040578/03.

DR N-PSDB; ABZ73639.

XX New human secreted proteins and nucleic acids, useful for detecting or

PT treating cancer or other hyperproliferative disorders, autoimmune

PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.

XX Claim 13; Page 1491; 2474pp; English.

PS ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted

XX protein genes, and ABP00947-ABP01363 represent the proteins they encode.

CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The

CC invention also encompasses antibodies specific for the secreted proteins,

CC the use of the secreted proteins in drug screening and recombinant

CC vectors and host cells comprising a nucleic acid of the invention. The

CC secreted proteins are thought to be involved in biological activities

CC associated with cellular signalling, cellular differentiation, cell

CC migration, prohormone activation and neurotransmitter activity. The

CC secreted proteins, nucleic acids encoding them, antibodies or antibody

CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification, chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein of the invention
XX
SQ Sequence 37 AA;
Query Match 22.3%; Score 49; DB 6; Length 37;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 26 VRRHVPVAVWL 36
Db 3 LRRHFPALWVL 13
RESULT 4
ADA98381
ID ADA98381 standard; protein; 37 AA.
XX
AC ADA98381;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein sequence #222.
XX
KW human; secreted protein; cardiovascular disorder; arrhythmia;
KW atherosclerosis; stroke; endocarditis; congestive heart failure;
KW rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;
KW migraine; thrombosis; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; renal disorder; proliferative disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003004623-A2.
XX
PD 16-JAN-2003.
XX
PF 26-MAR-2002; 2002WO-US009922.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-247946/24.
XX
PT New human secreted polypeptide and nucleic acid molecules, useful for
PT diagnosing, preventing, prognosticating or treating cardiovascular
PT disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
PT thrombosis).
XX
PS Claim 1; SEQ ID NO 490; 1572pp; English.
XX
CC The invention comprises the amino acid and coding sequence of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful in the treatment of cardiovascular disorders, such as: arrhythmia,
CC atherosclerosis, stroke, endocarditis, congestive heart failure,
CC rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins,
CC migraine, or thrombosis. The DNA and protein sequences may also be used
CC for treating or preventing: neural disorders, immune system disorders,
CC muscular disorders, reproductive disorders, gastrointestinal disorders,

CC pulmonary disorders, renal disorders, proliferative disorders and/or
CC cancerous diseases. The present amino acid sequence represents a human
CC secreted protein of the invention. NOTE: The present sequence is shown on
CC the WIPO website.
XX
SQ Sequence 37 AA;
Query Match 22.3%; Score 49; DB 6; Length 37;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 26 VRRHVPVAVWL 36
Db 3 LRRHFPALWVL 13
RESULT 5
ABO55642
ID ABO55642 standard; protein; 40 AA.
XX
AC ABO55642;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #1876.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 29276; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing

CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX
SQ Sequence 40 AA;

Query Match 22.0%; Score 48.5; DB 8; Length 40;
Best Local Similarity 55.0%; Pred. No. 29;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 12 WVPASPDGTGLDPLTVRRHVP 31
|: || | |||||: |
Db 18 WLFASAD-NLDPLTLRENTP 36

RESULT 6
AAG57081
ID AAG57081 standard; protein; 34 AA.

XX AAG57081;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 73511.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 20.5%; Score 45; DB 3; Length 34;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 7 QTEDSWVPASPDGTGLDPL 24
|: ||| | || | ||
Db 19 QSRDSWFP--FDLRLRPL 34

RESULT 7

AAM96628

ID AAM96628 standard; protein; 41 AA.

XX AC AAM96628;

XX XX

DT 21-NOV-2001 (first entry)

XX XX

DE Human reproductive system related antigen SEQ ID NO: 5286.

XX KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX OS Homo sapiens.

XX XX

PN WO200155320-A2.

XX XX

PD 02-AUG-2001.

XX XX

PF 17-JAN-2001; 2001WO-US001339.

XX XX

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-021680P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226279P.

KW chronic nephritis; blood-related disorder; thrombosis.

XX	OS	Homo sapiens.	PR	25-SEP-2000;	2000US-0234998P.
XX	OS		PR	26-SEP-2000;	2000US-0235484P.
XX	PN	WO200155316-A2.	PR	27-SEP-2000;	2000US-0235834P.
XX	PD		PR	29-SEP-2000;	2000US-0235836P.
XX	PD	02-AUG-2001.	PR	29-SEP-2000;	2000US-0236327P.
XX	PF		PR	29-SEP-2000;	2000US-0236367P.
XX	PF	17-JAN-2001;	PR	29-SEP-2000;	2000US-0236368P.
XX	XX		PR	29-SEP-2000;	2000US-0236369P.
PR	PR	31-JAN-2000;	PR	29-SEP-2000;	2000US-0236370P.
PR	PR	04-FEB-2000;	PR	02-OCT-2000;	2000US-0236802P.
PR	PR	24-FEB-2000;	PR	02-OCT-2000;	2000US-0237037P.
PR	PR	02-MAR-2000;	PR	02-OCT-2000;	2000US-0237038P.
PR	PR	16-MAR-2000;	PR	02-OCT-2000;	2000US-0237039P.
PR	PR	17-MAR-2000;	PR	02-OCT-2000;	2000US-0237040P.
PR	PR	18-APR-2000;	PR	13-OCT-2000;	2000US-0239935P.
PR	PR	19-MAY-2000;	PR	13-OCT-2000;	2000US-0239937P.
PR	PR	07-JUN-2000;	PR	20-OCT-2000;	2000US-0240960P.
PR	PR	28-JUN-2000;	PR	20-OCT-2000;	2000US-0241221P.
PR	PR	30-JUN-2000;	PR	20-OCT-2000;	2000US-0241785P.
PR	PR	07-JUL-2000;	PR	20-OCT-2000;	2000US-0241786P.
PR	PR	11-JUL-2000;	PR	20-OCT-2000;	2000US-0241787P.
PR	PR	14-JUL-2000;	PR	20-OCT-2000;	2000US-0241808P.
PR	PR	26-JUL-2000;	PR	20-OCT-2000;	2000US-0241809P.
PR	PR	26-JUL-2000;	PR	20-OCT-2000;	2000US-0241826P.
PR	PR	14-AUG-2000;	PR	01-NOV-2000;	2000US-0244617P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246474P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246475P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246476P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246477P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246478P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246523P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246524P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246525P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246526P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246527P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246528P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246532P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246609P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246610P.
PR	PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246611P.
PR	PR	18-AUG-2000;	PR	08-NOV-2000;	2000US-0246613P.
PR	PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249207P.
PR	PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249208P.
PR	PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249209P.
PR	PR	23-AUG-2000;	PR	17-NOV-2000;	2000US-0249210P.
PR	PR	30-AUG-2000;	PR	17-NOV-2000;	2000US-0249211P.
PR	PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249212P.
PR	PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249213P.
PR	PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249214P.
PR	PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249215P.
PR	PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249216P.
PR	PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249217P.
PR	PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249218P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249244P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249245P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249264P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249265P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249297P.
PR	PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249299P.
PR	PR	12-SEP-2000;	PR	17-NOV-2000;	2000US-0249300P.
PR	PR	14-SEP-2000;	PR	01-DEC-2000;	2000US-0250160P.
PR	PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0251030P.
PR	PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0251988P.
PR	PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0256719P.
PR	PR	14-SEP-2000;	PR	06-DEC-2000;	2000US-0251479P.
PR	PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251856P.
PR	PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251868P.
PR	PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251869P.
PR	PR	21-SEP-2000;	PR	08-DEC-2000;	2000US-0251989P.
PR	PR	21-SEP-2000;	PR	08-DEC-2000;	2000US-0251990P.
PR	PR	25-SEP-2000;	PR	11-DEC-2000;	2000US-0254097P.
			PR	05-JAN-2001;	2001US-0259678P.


```

XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Barash SC, Ruben SM;
XX PI WPI; 2001-451929/48.
DR DR N-PSDB; AAS40271.
XX PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the reproductive system including prostate cancer
PT and also for testing and detection e.g. diagnosis.
XX Claim 11; SEQ ID NO 423; 546pp; English.
XX The invention relates to novel isolated human prostate cancer antigen
CC polynucleotides (I) and polypeptides (II). (I) and (II) are useful for
CC preventing, treating or ameliorating a medical condition when
CC administered. (I), (II) and the antibody to (II) are useful for treating,
CC preventing and/ or prognosing disorders related to the reproductive
CC system including prostate cancers; urinary disorders e.g. chronic
CC nephritis; and blood-related disorders e.g. thrombosis. (II) can be used
CC for testing and detection e.g. as a chromosomal marker and in forensics.
CC (I) and the anti-(II) antibody can be used in testing and detection in
CC immunoassays. AAU22702-AAU22913 represent the human prostate cancer
CC antigen amino acid sequences, and related amino acid sequences of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences
XX Sequence 41 AA;
SQ Query Match 20.0%; Score 44; DB 4; Length 41;
Best Local Similarity 35.5%; Pred. No. 1.3e+02;
Matches 11; Conservative 7; Mismatches 9; Indels 4; Gaps 2
QY QY 2 KLRFLQTESWVPASPDGTGLDPLTVR---RH 29
    :|||: | :|:|| ||| :| :|
DB DB 2 QIRFVAIEANWIPACL-CGLDSHHIKSSTRH 31
RESULT 9
AAG66872
ID ID AAG66872 standard; peptide; 25 AA.
XX AC AAG66872;
XX DT 29-JAN-2002 (first entry)
XX DE Human ErbB2 peptide #3.
XX OS Homo sapiens.
XX PN JP2001218589-A.
XX PD 14-AUG-2001.
XX PF 28-APR-2000; 2000JP-00130787.
XX PR 02-DEC-1999; 99JP-00343587.
XX PA (TANB ) TT PHARM INC.
XX DR WPI; 2001-650745/75.
XX Novel peptide useful in drug research has an immune activating ability.
XX Example 1; Page 5; 14pp; Japanese.
XX The invention relates to 10 human ErbB2 (avian erythroblastic leukaemia
CC viral oncogene homologue 2) peptides with immunostimulatory activity. The

```

peptides of the invention, which are 15-25 amino acids in length, bind MHC (major histocompatibility complex) class II molecules on antigen-presenting cells and are capable of activating epitope-specific CD4+ helper T cells. The invention also relates to a cell-mediated immunity activation composition comprising an ErbB2 epitope peptide of the invention, nucleic acids encoding a peptide of the invention, vectors and transfectants comprising such nucleic acids, a method of determining the number of CD4+ T cells reactive with ErbB2, and a method of preparing an ErbB2 epitope peptide. The peptides of the invention are useful as immune system activators and in the field of drug research. The present sequence represents an ErbB2 peptide corresponding to residues 31-55 of human ErbB2 used in an exemplification of the invention

Sequence 25 AA;

Query Match 19.3%; Score 42.5; DB 4; Length 25;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:| ||| |||
Db 5 LPASPETHLDML---RHL 19

RESULT 10
AAW59346
ID AAW59346 standard; protein; 32 AA.
XX
AC AAW59346;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human ErbB2 extracellular domain protein fragment (aa 22-53).
XX
KW ErbB2; epitope; monoclonal antibody; identification; apoptosis;
KW detection; treatment; disorder; tumour; benign; malignant; leukaemia;
KW lymphoid malignancy; inflammation; angiogenic; immunological.
XX
OS Homo sapiens.
XX
PN WO9817797-A1.
XX
PD 30-APR-1998.
XX
PF 09-OCT-1997; 97WO-US018385.
XX
PR 18-OCT-1996; 96US-00731794.
XX
PA (GETH) GENENTECH INC.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Fendly BM, Phillips GD, Scheuermann RH, Uhr JW;
XX
DR WPI; 1998-261493/23.
XX
PT New anti-ErbB2 antibodies - which induce apoptosis in cells which
PT overexpress ErbB2, used to treat e.g. tumours or inflammatory, angiogenic
PT or immunologic disorders.
XX
PS Disclosure; Page 45; 72pp; English.
XX

This sequence represents a fragment of the human ErbB2 extracellular domain corresponding to amino acids 22-53 which is capable of binding the monoclonal antibodies 7C2 and 7F3. This region is used in a method to identify antibodies which bind to ErbB2 and induce apoptosis of a cell which overexpresses ErbB2. The antibodies can also be used to detect ErbB2 on a cell. The antibodies can be used for treating disorders such as benign or malignant tumours, (e.g. renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukemia's and lymphoid malignancies, other disorders such as neuronal, glial, astrocytal, hypothalamic and other disorders macrophagal, epithelial, stromal and blastocoeleic disorders and

```
CC inflammatory, angiogenic and immunologic disorders
XX
SQ Sequence 32 AA;

Query Match      19.3%; Score 42.5; DB 2; Length 32;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY      13 VPASPDGTGLDPLTVRRHV 30
Db      14 LPASPETHLDMML---RHL 28
      :||||:|||||
RESULT 11
AAY14575
ID AAY14575 standard; peptide; 32 AA.
XX
AC AAY14575;
XX
DT 14-SEP-1999. (first entry)
XX
DE Human erbB2 receptor-derived epitope 7C2/7F3.
XX
KW Domain; human; erbB2; receptor; antibody; overexpression; anthracycline;
KW chemotherapeutic agent; tumour; cancer; breast; squamous cell; cervical;
KW non-small-cell lung; gastrointestinal; pancreatic; glioblastoma; ovarian;
KW liver; bladder; hepatoma; colon; colorectal; endometrial carcinoma;
KW salivary gland; kidney; prostate; vulva; thyroid; hepatic carcinoma;
KW head-and-neck; epitope.
XX
OS Homo sapiens.
XX
PN WO9931140-A1.
XX
PD 24-JUN-1999.
XX
PF 10-DEC-1998; 98WO-US026266.
XX
PR 12-DEC-1997; 97US-0069346P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Shak S, Paton VE;
XX
DR WPI; 1999-395164/33.
XX
PT New combination of an anti-ErbB2 antibody and a chemotherapeutic agent.
XX
PS Disclosure; Page 35; 42pp; English.
XX
CC This sequence represents the 7C2/7F3 epitope from domain 1 of the human
CC erbB2 receptor protein. ErbB2 protein is used for raising antibodies for
CC treating a human with a disorder characterised by overexpression of the
CC ErbB2 receptor, using a combination of an anti-ErbB2 antibody and a
CC chemotherapeutic agent other than an anthracycline derivative. Disorders
CC characterised by overexpression of the ErbB2 receptor include benign or
CC malignant tumours. The cancer may be breast, squamous cell, (non)-small-
CC cell lung, gastrointestinal, pancreatic, glioblastoma, cervical, ovarian,
CC liver, bladder, hepatoma, colon, colorectal, endometrial carcinoma,
CC salivary gland carcinoma, kidney, prostate, vulval or thyroid cancer,
CC hepatic carcinoma and various forms of head-and-neck cancer
XX
SQ Sequence 32 AA;

Query Match      19.3%; Score 42.5; DB 2; Length 32;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY      13 VPASPDGTGLDPLTVRRHV 30
Db      14 LPASPETHLDMML---RHL 28
      :||||:|||||
RESULT 12
AAB62075
ID AAB62075 standard; peptide; 32 AA.
XX
AC AAB62075;
XX
DT 29-MAY-2001 (first entry)
XX
DE ErbB2 7C2/7f3 epitope.
XX
KW ErbB2 receptor; anti-ErbB2 antibody; tumour; cancer; breast cancer;
KW leukemia; cytostatic; carcinoma; epitope.
XX
OS Homo sapiens.
XX
PN WO200115730-A1.
XX
PD 08-MAR-2001.
XX
PF 25-AUG-2000; 2000WO-US023391.
XX
PR 27-AUG-1999; 99US-0151018P.
PR 23-JUN-2000; 2000US-0213822P.
XX
(GETH ) GENENTECH INC.
XX
Baughman SA, Shak S;
XX
WPI; 2001-235058/24.
XX
PT Treating a human patient with a disorder characterized by overexpression
PT of ErbB2 receptor such as a tumor or cancer (e.g. metastatic breast
PT carcinoma) comprises administering several dosages of an anti-ErbB2
PT antibody.
XX
PS Disclosure; Page 64; 71pp; English.
XX
CC The invention relates to treating cancer in a human susceptible to or
CC diagnosed with a disorder characterized by overexpression of ErbB2
CC receptor. The method comprises administering an initial dose of at least
CC approx.5 mg/kg of the anti-ErbB2 antibody, and administering subsequent
CC doses of the antibody in approximately the same amount or less than the
CC initial dose. The method is useful for treating a disorder characterized
CC by overexpression of ErbB2 receptor, particularly a benign or malignant
CC tumour, or a cancer. The cancer is selected from breast cancer
CC (particularly metastatic breast carcinoma), leukemia, squamous cell
CC cancer, small-cell lung cancer, non-small cell lung cancer, pancreatic
CC cancer, gastrointestinal cancer, glioblastoma, cervical cancer, ovarian
CC cancer, liver cancer, bladder cancer, hepatoma, colon cancer, colorectal
CC cancer, endometrial carcinoma, salivary gland carcinoma, kidney cancer,
CC liver cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic
CC carcinoma and various types of head and neck cancer. The present sequence
CC represents the 7C2/7F3 peptide epitope of ErbB2, recognised by MAbs 7C2
CC and 7F3 as determined by deletion mapping
XX
SQ Sequence 32 AA;

Query Match      19.3%; Score 42.5; DB 4; Length 32;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY      13 VPASPDGTGLDPLTVRRHV 30
Db      14 LPASPETHLDMML---RHL 28
      :||||:|||||
RESULT 13
AAB48764
ID AAB48764 standard; protein; 32 AA.
XX
AC AAB48764;
XX
DT 09-MAR-2001 (first entry)
```

XX DE Human ErbB2 7C2/7F3 epitope (residues 22-53).

XX KW Human ErbB2; HER2; c-ErbB-2; monoclonal antibody 7C2; 7F3; epitope;

KW extracellular domain; humanised anti-ErbB2 antibody; antiproliferative;

KW cancer; tumour; breast; squamous cell tumour; small cell lung tumour;

KW non-small cell lung tumour; gastrointestinal tumour; pancreatic tumour;

KW colorectal tumour; glioblastoma; hepatoma; cervix; ovary; the liver;

KW bladder; colon; colorectal; endometrium; vulva; salivary gland; kidney;

KW prostate; head; neck; tumour antigen; anticancer.

XX OS Homo sapiens.

XX WO200069460-A1.

XX PN 23-NOV-2000.

XX PD 09-MAY-2000; 2000WO-US012552.

XX PF 14-MAY-1999; 99US-0134085P.

XX PR (GETH) GENENTECH INC.

XX PA Cohen RL;

XX PI WPI; 2001-016161/02.

XX DR Treating a patient susceptible to, or diagnosed with, a tumor associated

XX PT with breast, squamous cell, colon, kidney, liver by administering anti-

XX PT ErbB2 antibody before and after surgical removal of the tumor.

XX PS Disclosure; Page 34-35; 39pp; English.

XX CC The invention relates to a novel method of treating a human patient

CC susceptible to or diagnosed with a tumour expressing the ErbB2 (HER2, c-

CC Erb-B2) tumour antigen. The method comprises treating the patient with an

CC anti-ErbB2 antibody, removing the tumour surgically, and then treating

CC the patient with the anti-ErbB2 antibody or a chemotherapeutic agent. In

CC particular, the antibody is a humanised version of the murine anti-ErbB2

CC antibody 4D5, which binds to the an epitope on the extracellular domain

CC of ErbB2 (residues 561-625; AAB48766). The method of the invention is

CC used to treat a human patient susceptible to, or diagnosed with breast

CC tumour, squamous cell tumour, small cell lung tumour, non-small cell lung

CC tumour, gastrointestinal tumour, pancreatic tumour, colorectal tumour,

CC glioblastoma, hepatoma, cervical tumour, ovarian tumour, or a tumour of

CC the liver, bladder, colon, endometrium, vulva, salivary gland, kidney,

CC prostate, head, or neck. The present sequence represents the epitope of

CC the human ErbB2 (residues 22-58) which binds to the monoclonal antibodies

CC 7C2 and 7F3. Humanised versions of these antibodies may also be used in

XX CC the method of the invention

XX SQ Sequence 32 AA;

Query Match 19.3%; Score 42.5; DB 4; Length 32;

Best Local Similarity 55.6%; Pred. NO. 1.6e+02;

Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDITGLDPLTVRRHV 30

Db :|||:| | | | |

14 LPASPETHLDML---RHL 28

RESULT 14

ADB84514

ID ADB84514 standard; protein; 27 AA.

XX AC ADB84514;

XX DT 04-DEC-2003 (first entry)

XX DE MSRV-1B pol region isolate #3 protein #3.

XX KW MSRV; multiple sclerosis; rheumatoid arthritis; gag; pol;

KW reverse transcriptase; ribonuclease H; antigen.

XX Multiple sclerosis associated retrovirus.

OS US2003039664-A1.

XX 27-FEB-2003.

PD 26-NOV-1997; 97US-00979847.

XX 26-NOV-1996; 96US-00756429.

XX (PERR/) PERRON H.

PA (BESE/) BESEME F.

PA (BEDI/) BEDIN F.

PA (PARA/) PARANHOS-BACCALA G.

PA (KOMU/) KOMURIAN-PRADEL F.

PA (JOLI/) JOLIVET-REYNAUD C.

PA (MAND/) MANDRAND B.

PA (GARS/) GARSON J A.

PA (TUKA/) TUKE P W.

XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;

PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;

PI WPI; 2003-512253/48.

XX N-PSDB; ADB84553.

DR New isolated or purified nucleic acid associated with multiple sclerosis

XX and/or rheumatoid arthritis, useful for detecting a virus associated with

PT multiple sclerosis or rheumatoid arthritis in a biological sample.

PS Example 1; Fig 2; 193pp; English.

XX The invention relates to an isolated or purified nucleic acid from a

CC virus associated with multiple sclerosis and/or rheumatoid arthritis,

CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise

CC pol, gag or reverse transcriptase genes (or their fragments) encoding the

CC proteins or defined peptides (including immunodominant peptides,

CC antigenic peptides or conserved motifs). Also included are a process for

CC detecting a virus associated with multiple sclerosis or rheumatoid

CC arthritis in a biological sample, a nucleic acid probe for the detection

CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a

CC primer for the amplification by polymerisation of a nucleic acid of a

CC viral material associated with multiple sclerosis or rheumatoid

CC arthritis, a polypeptide exhibiting an inhibitory activity on the

CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,

CC and an antibody directed against the MSRV-1 virus obtained by

CC immunologically reacting a human or animal body or cells with an

CC immunogenic agent consisting of the antigenic polypeptide defined above.

CC The nucleic acids are useful for detecting a biological sample a virus

CC associated with multiple sclerosis or rheumatoid arthritis, or for

CC detecting in a biological sample, the presence of or exposure to a virus

CC associated with multiple sclerosis or rheumatoid arthritis. The present

CC sequence is an MSRV protein, protein fragment, or consensus sequence.

CC Note: The SEQ ID numbers for the sequences as displayed in the main body

CC of the patent do not match the SEQ ID numbers in the sequence listing.

CC Consequently those sequences mentioned in the claims may not be the

XX sequences the authors intended to claim.

SQ Sequence 27 AA;

Query Match 19.1%; Score 42; DB 7; Length 27;

Best Local Similarity 34.6%; Pred. No. 1.5e+02;

Matches 9; Conservative 4; Mismatches 7; Indels 6; Gaps 1;

QY 19 TGLDPLTVRRHVPA-----VWVLLS 38

Db :|:|:| | | | |

1 SGIAPYILARHXPKTXVNSHTWLLS 26

RESULT 15

AAY98932

Search completed: January 26, 2005, 15:49:52
Job time : 163 secs

ID AAY98932 standard; peptide; 15 AA.
XX
AC AAY98932;
XX
DT 07-AUG-2000 (first entry)
XX
DE HLA class II binding antigen epitope peptide #121.
XX
KW Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
KW immune response; chronic viral disease; cancer; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW allograft rejection; allergy; lyme disease; hepatitis; prostate cancer;
KW glomerulonephritis; food hypersensitivity; malaria.
XX
OS Unidentified.
XX
PN WO9961916-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US012066.
XX
PR 29-MAY-1998; 98US-0087192P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Southwood S, Sidney J;
XX
DR WPI; 2000-097143/08.
XX
PT New compositions containing immunogenic peptide epitopes for various HLA
PT class II DR molecules useful for inducing helper T cell response.
XX
PS Claim 1; Page 42; 60pp; English.
XX
CC The present invention relates to a new pharmaceutical composition
CC comprising a unit dose form of a peptide, or analogue, comprising an
CC epitope selected from those represented by peptides AAY98812-Y99339 which
CC are derived from various antigens for various human leucocyte antigen
CC class DR molecules, representative of the world wide population. The
CC peptide/analogue binds to an HLA class II molecule at an IC-50 of less
CC than or equal to 1,000 nM. The pharmaceutical can be used to induce a
CC helper T cell response. The pharmaceutical focuses the immune response
CC towards selected determinants and could therefore be used in cases of
CC chronic viral diseases and cancer. Examples of diseases that can be
CC treated using the peptide containing pharmaceutical include autoimmune
CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
CC gravis), allograft rejection, allergies, lyme disease, hepatitis, post-
CC streptococcal endocarditis or glomerulonephritis and food
CC hypersensitivities. The peptide epitopes can be used to enhance immune
CC responses against other immunogens administered with the peptides.
CC Diseases which can be treated using immunogenic mixtures include prostate
CC cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
CC carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
CC used to make monoclonal antibodies useful as potential diagnostic or
CC therapeutic agents. The peptides may also be useful as diagnostic
CC reagents, for example, to determine the susceptibility of an individual
CC to a treatment regimen. Also, the peptides may be used to predict which
CC individuals will be at substantial risk of developing chronic infection.
CC The selection of appropriate T and B cell epitopes should allow the
CC development of epitope based vaccines particularly towards conserved
CC epitopes of pathogens which are characterized by high sequence
CC variability such as HIV, HCV and Malaria
XX

SQ Sequence 15 AA;

Query Match 18.9%; Score 41.5; DB 3; Length 15;
Best Local Similarity 58.8%; Pred. NO. 86;
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRH 29
:||||:||||
Db 2 LPASPETHLDML---RH 15

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:53:27 ; Search time 144 Seconds
(without alignments)
102.867 Million cell updates/sec

Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOTEDSWVPASPDG.....DPLTVRRHVPVAVVLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 391627

Minimum DB seq length: 0
Maximum DB seq length: 41

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	22.0	40	14	US-10-029-386-29276 Sequence 29276, A
2	44	20.0	41	10	US-09-764-891-5286 Sequence 5286, Ap
3	44	20.0	41	14	US-10-091-572-423 Sequence 423, App
4	43	19.5	41	15	US-10-424-599-213057 Sequence 213057,
5	42.5	19.3	32	14	US-10-356-824-2 Sequence 2, Appli
6	42.5	19.3	32	14	US-10-406-925-2 Sequence 2, Appli
7	42.5	19.3	32	14	US-10-429-519-2 Sequence 2, Appli
8	42.5	19.3	32	15	US-10-356-824-2 Sequence 2, Appli
9	42.5	19.3	32	15	US-10-600-152-2 Sequence 2, Appli
10	42.5	19.3	32	17	US-10-909-998-2 Sequence 2, Appli
11	42	19.1	37	17	US-10-425-115-334035 Sequence 334035,
12	42	19.1	40	17	US-10-425-115-249550 Sequence 249550,
13	41.5	18.9	15	15	US-10-149-138-3951 Sequence 3951, Ap

14	41.5	18.9	15	15	US-10-149-138-4508	Sequence 4508, Ap
15	41.5	18.9	15	16	US-10-149-138-3951	Sequence 3951, Ap
16	41.5	18.9	15	16	US-10-149-138-4508	Sequence 4508, Ap
17	41	18.6	25	9	US-09-729-835-96	Sequence 96, Appl
18	41	18.6	25	15	US-10-373-809-96	Sequence 96, Appl
19	40.5	18.4	32	14	US-10-080-254-73	Sequence 73, Appl
20	40.5	18.4	32	14	US-10-242-355-617	Sequence 617, App
21	40.5	18.4	40	15	US-10-424-599-187800	Sequence 187800,
22	40	18.2	37	17	US-10-425-115-228363	Sequence 228363,
23	40	18.2	38	9	US-09-864-761-42375	Sequence 42375, A
24	40	18.2	38	9	US-09-843-676-27	Sequence 27, Appl
25	40	18.2	38	9	US-09-766-253-27	Sequence 27, Appl
26	40	18.2	38	10	US-09-438-486-27	Sequence 27, Appl
27	40	18.2	38	14	US-10-053-758-27	Sequence 27, Appl
28	40	18.2	38	14	US-10-054-295-27	Sequence 27, Appl
29	40	18.2	38	14	US-10-054-611-27	Sequence 27, Appl
30	40	18.2	38	14	US-10-325-810-217	Sequence 217, App
31	40	18.2	38	17	US-10-877-124-217	Sequence 217, App
32	40	18.2	38	17	US-10-877-022-217	Sequence 217, App
33	40	18.2	40	10	US-09-764-891-3493	Sequence 3493, Ap
34	39	17.7	11	15	US-10-149-138-570	Sequence 570, App
35	39	17.7	11	15	US-10-149-138-1921	Sequence 1921, Ap
36	39	17.7	11	16	US-10-149-138-570	Sequence 570, App
37	39	17.7	11	16	US-10-149-138-1921	Sequence 1921, Ap
38	39	17.7	33	12	US-09-855-604-715	Sequence 715, App
39	39	17.7	38	17	US-10-425-115-308261	Sequence 308261,
40	39	17.7	41	10	US-09-826-734-106	Sequence 106, App
41	38	17.3	11	15	US-10-149-138-1515	Sequence 1515, Ap
42	38	17.3	11	15	US-10-149-138-2164	Sequence 2164, Ap
43	38	17.3	11	16	US-10-149-138-1515	Sequence 1515, Ap
44	38	17.3	11	16	US-10-149-138-2164	Sequence 2164, Ap
45	38	17.3	15	15	US-10-149-138-3741	Sequence 3741, Ap

ALIGNMENTS

RESULT 1

US-10-029-386-29276
; Sequence 29276, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29276
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
US-10-029-386-29276

Query Match 22.0%; Score 48.5; DB 14; Length 40;
Best Local Similarity 55.0%; Pred. No. 55;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;
QY 12 WVPASPDGTGLDPLTVRRHVP 31
|: ||| | |||||: |
Db 18 WLFASAD-NLDPLTLRPNTP 36

RESULT 4

US-10-424-399-213057
; Sequence 213057, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

APPLICANT: Cao Yongwei

; SEQ ID NO 213057

; TYPE: PRT

; **FEATURE:**

Query Match

Matches 12; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 5

; Sequence 2, Application US/10356824

GENERAL INFORMATION:

APPLICANT: Paton, Virginia

; FILE REFERENCE: P1256R1

; CURRENT FILING DATE: 2003-02-03

;; PRIOR FILING DATE: 1998-12-10

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12

SEQ ID NO 2

; LENGTH: 32

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-356-824-2

Query Match 19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 6

US-10-406-925-2
; Sequence 2, Application US/10406925
; Publication No. US20030170234A1
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1256R3

; CURRENT APPLICATION NUMBER: US/10/406,925
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/209,023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9

; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-406-925-2

Query Match 19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 7

US-10-429-519-2
; Sequence 2, Application US/10429519
; Publication No. US20030170235A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1757R1

; CURRENT APPLICATION NUMBER: US/10/429,519
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/568,322
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US 60/134,085
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 4

; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-429-519-2

Query Match 19.3%; Score 42.5; DB 14; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 8
US-10-356-824-2
; Sequence 2, Application US/10356824
; Publication No. US20040037823A9
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1256R1

; CURRENT APPLICATION NUMBER: US/10/356,824
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/208,649
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9

; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-356-824-2

Query Match 19.3%; Score 42.5; DB 15; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 9

US-10-600-152-2
; Sequence 2, Application US/10600152
; Publication No. US20040037824A1
; GENERAL INFORMATION:
; APPLICANT: Baughman, Sharon A.
; APPLICANT: Shak Steven

; TITLE OF INVENTION: Dosages for Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1775R1
; CURRENT APPLICATION NUMBER: US/10/600,152
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: 09/648,067
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,018
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/213,822
; PRIOR FILING DATE: 2000-06-23

; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-600-152-2

Query Match 19.3%; Score 42.5; DB 15; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 10

US-10-909-998-2
; Sequence 2, Application US/10909998
; Publication No. US20050002928A1
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia

; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES

;
; FILE REFERENCE: P1256R3
; CURRENT APPLICATION NUMBER: US/10/909,998
; PRIOR FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/209,023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-909-998-2

Query Match 19.3%; Score 42.5; DB 17; Length 32;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRHV 30
:||||:|||||
Db 14 LPASPETHLDML---RHL 28

RESULT 11
US-10-425-115-334035
; Sequence 334035, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 334035
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_67752C.1.pep
US-10-425-115-334035

Query Match 19.1%; Score 42; DB 17; Length 37;
Best Local Similarity 53.8%; Pred. No. 3.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LRFLQTEDSWVPA 15
:|||:|||||
Db 25 IRLLETPDAWVTA 37

RESULT 12
US-10-425-115-249550
; Sequence 249550, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 249550
; LENGTH: 40
; TYPE: PRT

;
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_159174C.1.pep
US-10-425-115-249550

Query Match 19.1%; Score 42; DB 17; Length 40;
Best Local Similarity 36.4%; Pred. No. 4e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 9 EDSWVPASPDGTGLDPLTVRRHV 30
::||:|||||
Db 6 QETWM--DPAIGLTPYDVHKHV 25

RESULT 13
US-10-149-138-3951
; Sequence 3951, Application US/10149138
; Publication No. US20040018971A1
; GENERAL INFORMATION:
; APPLICANT: Fikes, John
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Bateban
; APPLICANT: Keogh, Elissa
; TITLE OF INVENTION: Inducing Cellular Immune Responses to
; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0140001
; CURRENT APPLICATION NUMBER: US/10/149,138
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/US00/33591
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/458,299
; PRIOR FILING DATE: 1999-12-11
; NUMBER OF SEQ ID NOS: 4641
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3951
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially Synthesized Peptide
US-10-149-138-3951

Query Match 18.9%; Score 41.5; DB 15; Length 15;
Best Local Similarity 58.8%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRH 29
:||||:|||||
Db 2 LPASPETHLDML---RH 15

RESULT 14
US-10-149-138-4508
; Sequence 4508, Application US/10149138
; Publication No. US20040018971A1
; GENERAL INFORMATION:
; APPLICANT: Fikes, John
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Keogh, Elissa
; TITLE OF INVENTION: Inducing Cellular Immune Responses to
; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0140001
; CURRENT APPLICATION NUMBER: US/10/149,138
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/US00/33591
; PRIOR FILING DATE: 2000-12-11

; PRIOR APPLICATION NUMBER: US 09/458,299
; PRIOR FILING DATE: 1999-12-11
; NUMBER OF SEQ ID NOS: 4641
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4508
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Peptide
US-10-149-138-4508

Query Match 18.9%; Score 41.5; DB 15; Length 15;
Best Local Similarity 58.8%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRH 29
:||||:||||||
Db 2 LPASPETHLDML---RH 15

RESULT 15
US-10-149-138-3951
; Sequence 3951, Application US/10149138
; Publication No. US20040121946A9
; GENERAL INFORMATION:
; APPLICANT: Fikes, John
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Keogh, Elissa
; TITLE OF INVENTION: Inducing Cellular Immune Responses to
; FILE REFERENCE: 2060.0140001
; CURRENT APPLICATION NUMBER: US/10/149,138
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/US00/33591
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/458,299
; PRIOR FILING DATE: 1999-12-11
; NUMBER OF SEQ ID NOS: 4641
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3951
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially Synthesized Peptide
US-10-149-138-3951

Query Match 18.9%; Score 41.5; DB 16; Length 15;
Best Local Similarity 58.8%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 13 VPASPDGTGLDPLTVRRH 29
:||||:||||||
Db 2 LPASPETHLDML---RH 15

Search completed: January 26, 2005, 16:05:52
Job time : 145 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 15:47:17 ; Search time 38 Seconds
(without alignments)
71.554 Million cell updates/sec

Title: US-10-057-813-14_COPY_85_125
Perfect score: 220
Sequence: 1 TKLRFLOTEDSWVPASPDGTG.....DPLTVRRHVPVAVWVLLSRDP 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 251498

Minimum DB seq length: 0
Maximum DB seq length: 41

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45.5	20.7	40	4	US-09-270-767-41398
2	45.5	20.7	40	4	US-09-270-767-56614
3	45	20.5	17	3	US-09-174-465D-6
4	45	20.5	17	3	US-09-599-564A-6
5	42.5	19.3	32	4	US-09-648-067A-2
6	41	18.6	25	4	US-09-257-179-96
7	40	18.2	38	3	US-08-851-843A-27
8	40	18.2	38	3	US-08-974-549A-217
9	40	18.2	38	3	US-08-854-050-27
10	40	18.2	38	3	US-09-430-323-27
11	40	18.2	38	4	US-09-402-181B-217
12	40	18.2	38	4	US-09-721-456-217
13	39	17.7	35	2	US-08-726-306A-39
14	39	17.7	38	4	US-09-270-767-56715
15	38.5	17.5	28	1	US-07-596-081A-14
16	38.5	17.5	33	1	US-08-086-428B-145
17	38.5	17.5	33	2	US-08-468-570-145
18	38.5	17.5	33	2	US-08-290-665A-249
19	38.5	17.5	33	4	US-08-466-601A-145
20	38.5	17.5	33	5	PCT-US95-10398-249
21	37.5	17.0	34	4	US-09-645-470-3
22	37.5	17.0	34	4	US-09-355-040-20
23	37	16.8	25	3	US-08-612-973-55
24	37	16.8	25	3	US-08-927-597-55
25	37	16.8	39	3	US-08-851-843A-26
26	37	16.8	39	3	US-08-974-549A-216
27	37	16.8	39	3	US-08-854-050-26

28	37	16.8	39	3	US-09-430-323-26	Sequence 26, Appl
29	37	16.8	39	4	US-09-402-181B-216	Sequence 216, App
30	37	16.8	39	4	US-09-721-456-216	Sequence 216, App
31	37	16.8	41	3	US-09-177-249-256	Sequence 256, App
32	36.5	16.6	34	1	US-08-324-977-26	Sequence 26, Appl
33	36.5	16.6	34	2	US-08-384-616-26	Sequence 26, Appl
34	36.5	16.6	34	2	US-08-904-686A-26	Sequence 26, Appl
35	36.5	16.6	34	3	US-09-315-850-26	Sequence 26, Appl
36	36	16.4	14	4	US-09-563-222C-171	Sequence 171, App
37	36	16.4	38	1	US-08-176-500-70	Sequence 70, Appl
38	36	16.4	38	1	US-08-471-052A-70	Sequence 70, Appl
39	36	16.4	38	1	US-08-189-331-70	Sequence 70, Appl
40	36	16.4	38	2	US-08-471-939-70	Sequence 70, Appl
41	36	16.4	38	2	US-08-471-800-70	Sequence 70, Appl
42	36	16.4	38	2	US-08-471-068-70	Sequence 70, Appl
43	35	15.9	11	3	US-09-113-977C-16	Sequence 16, Appl
44	35	15.9	11	4	US-09-351-048A-16	Sequence 16, Appl
45	35	15.9	37	4	US-09-461-325-242	Sequence 242, App

ALIGNMENTS

RESULT 1
US-09-270-767-41398
; Sequence 41398, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41398
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-41398

Query Match 20.7%; Score 45.5; DB 4; Length 40;
Best Local Similarity 36.4%; Pred. No. 19;
Matches 12; Conservative 4; Mismatches 14; Indels 3; Gaps 2;

QY 9 EDSWVPASPDGTGLDPLTVRRHVPVAVWVLLSRDP 41
Db 10 DSTWAPSIYTTG-SPLRRR--PLIWDLSPTTP 39

RESULT 2
US-09-270-767-56614
; Sequence 56614, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56614
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56614

Query Match 20.7%; Score 45.5; DB 4; Length 40;
Best Local Similarity 36.4%; Pred. No. 19;
Matches 12; Conservative 4; Mismatches 14; Indels 3; Gaps 2;

QY 9 EDSWVPASPDGTGLDPLTVRRHVPVAVWVLLSRDP 41

RESULT 7

US-08-851-843A-27
; Sequence 27, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-851-843A-27

Query Match 18.2%; Score 40; DB 3; Length 38;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;

Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RELOQT---EDSWVPAS 16

Db 18 RELETTAEKNDGWVPIS 34

RESULT 8

US-08-974-549A-217
; Sequence 217, Application US/08974549A
; Patent No. 6166178

; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-549A-217

Query Match

18.2%; Score 40; DB 3; Length 38;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT----EDSWVPAS 16
| | | | |
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 9
US-08-854-050-27
; Sequence 27, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-27

Query Match 18.2%; Score 40; DB 3; Length 38;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT----EDSWVPAS 16
| | | | |
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 10
US-09-430-323-27
; Sequence 27, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-430-323-27

Query Match 18.2%; Score 40; DB 3; Length 38;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 4 RFLQT----EDSWVPAS 16
| | | | |
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 11
US-09-402-181B-217
; Sequence 217, Application US/09402181B
; Patent No. 6610839
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 633
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,181B
; FILING DATE: 29-Sep-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ausenhus, Scott L.
; REGISTRATION NUMBER: 42,271
; REFERENCE/DOCKET NUMBER: 015389-002620US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-402-181B-217

Query Match 18.2%; Score 40; DB 4; Length 38;
Best Local Similarity 52.9%; Pred. NO. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
QY 4 RFLQT----EDSWVPAS 16
|||:| |||
Db 18 RFLRTAEKNDGWVPIS 34

RESULT 12
US-09-721-456-217
; Sequence 217, Application US/09721456
; Patent No. 6617110
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/721,456
; FILING DATE: 22-No. 6617110-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-721-456-217

Query Match 18.2%; Score 40; DB 4; Length 38;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
QY 4 RFLQT-----EDSWVPAS 16
| | | | | | | | | |
Db 18 RFLRTTAEKNDGWVPIS 34

RESULT 13
US-08-726-306A-39
; Sequence 39, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-726-306A-39

Query Match 17.7%; Score 39; DB 2; Length 35;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TKLRFLOQTEDSWVP 14
| | | | | | | | | |
Db 22 TKLFFTHTRHSWNP 35

RESULT 14
US-09-270-767-56715
; Sequence 56715, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56715
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56715

Query Match 17.7%; Score 39; DB 4; Length 38;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 12 WVPASPT 19
| : | : | | |
Db 11 WIPETPT 18

RESULT 15
US-07-596-081A-14
; Sequence 14, Application US/07596081A
; Patent No. 5194586
; GENERAL INFORMATION:
; APPLICANT: Maeda, Yoshiaki
; APPLICANT: Shiraki, Hiroshi
; APPLICANT: Washitani, Yukiko
; APPLICANT: Kuroda, Naotaka
; APPLICANT: Yamada, Kyoko
; APPLICANT: Oka, Kiichiro
; APPLICANT: Namba, Toshihiro
; TITLE OF INVENTION: Peptides and Use Thereof
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eric S. Spector
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/596,081A
; FILING DATE: 19901011
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 266983/1989
; FILING DATE: 13-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Spector, Eric S.
; REGISTRATION NUMBER: 22495
; REFERENCE/DOCKET NUMBER: 513904
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-596-081A-14

Query Match 17.5%; Score 38.5; DB 1; Length 28;
Best Local Similarity 38.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

QY 2 KLRFLQTEDSWVPAS-----PDTGLD 22
| : | | | | | : | : | |
Db 1 KKKFLNTEPSQLPPTAPPLPHSNLD 26

Search completed: January 26, 2005, 16:02:48
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 31, 2005, 18:09:35 ; Search time 1821 Seconds
(without alignments)
7089.560 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaagctacgtttccttcag 273

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 3273528

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29.2	10.7	166	11	BV165390	BV165390 MsU63-12
C 2	28.8	10.5	133	4	AB126591	AB126591 Canis fam
C 3	28.8	10.5	201	11	BV200943	BV200943 sqnm20546
C 4	28.8	10.5	227	9	HSA506605	AJ506605 Homo sapi
C 5	28.4	10.4	262	6	CQ660987	CQ660987 Sequence
C 6	28.2	10.3	273	11	MMSTS122	Z36745 M.musculus
7	28	10.3	273	10	MUSNR2C01	L35014 Mouse N-met
C 8	27.8	10.2	154	6	CQ679231	CQ679231 Sequence
9	27.8	10.2	194	6	CQ443803	CQ443803 Sequence
C 10	27.4	10.0	240	9	HSA230371	AJ230371 Homo sapi
C 11	27.2	10.0	99	6	CQ734281	CQ734281 Sequence
C 12	27.2	10.0	235	6	AX798801	AX798801 Sequence
C 13	27	9.9	160	11	RICG332A	D25339 Oryza sativ
14	27	9.9	241	9	HS140C7F	Z59291 H.sapiens C
15	27	9.9	264	6	BD237006	BD237006 Compounds
16	27	9.9	264	6	AR225406	AR225406 Sequence
17	27	9.9	264	6	AX321476	AX321476 Sequence
18	27	9.9	264	9	S68587	S68587 platelet-ty
19	27	9.9	264	9	S68588	S68588 platelet-ty

C 20	26.8	9.8	208	11	HSA310747	AJ310747 Homo sapi
21	26.8	9.8	241	11	HSPE07C12	AL009365 H.sapiens
C 22	26.8	9.8	251	11	BV167732	BV167732 sqnm6347
C 23	26.8	9.8	269	9	HSDNASAAAT	Z22256 H.sapiens D
C 24	26.6	9.7	245	5	AB035828	AB035828 Coturnix
25	26.4	9.7	167	6	CQ080407	CQ080407 Sequence
26	26.4	9.7	167	6	CQ114259	CQ114259 Sequence
27	26.4	9.7	167	6	CQ153138	CQ153138 Sequence
28	26.4	9.7	167	6	CQ186132	CQ186132 Sequence
29	26.4	9.7	167	6	CQ236438	CQ236438 Sequence
30	26.4	9.7	167	6	CQ274027	CQ274027 Sequence
31	26.4	9.7	167	6	CQ311246	CQ311246 Sequence
32	26.4	9.7	167	6	CQ348333	CQ348333 Sequence
C 33	26.4	9.7	198	6	CQ081138	CQ081138 Sequence
C 34	26.4	9.7	198	6	CQ115663	CQ115663 Sequence
C 35	26.4	9.7	198	6	CQ154439	CQ154439 Sequence
C 36	26.4	9.7	198	6	CQ186999	CQ186999 Sequence
C 37	26.4	9.7	198	6	CQ237699	CQ237699 Sequence
C 38	26.4	9.7	198	6	CQ275317	CQ275317 Sequence
C 39	26.4	9.7	198	6	CQ312286	CQ312286 Sequence
C 40	26.4	9.7	198	6	CQ349715	CQ349715 Sequence
C 41	26.2	9.6	240	9	HSA415844	AJ415844 Homo sapi
C 42	26	9.5	72	10	AF093879	AF093879 Mus muscu
C 43	26	9.5	92	6	AR204954	AR204954 Sequence
44	26	9.5	166	11	BV165391	BV165391 MsU63-13
C 45	26	9.5	171	3	AF135376	AF135376 Semibalan

ALIGNMENTS

RESULT 1
BV165390
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BV165390 166 bp DNA linear STS 02-AUG-2004
MsU63-12 PCR fragment of the molecular marker, MsU63-Vr-TC1966
Vigna radiata STS genomic, sequence tagged site.
BV165390
BV165390.1 GI:50880891
STS.
Vigna radiata
Vigna radiata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Vigna.
1 (bases 1 to 166)
Choi,H.K., Doyle,J. and Cook,D.R.
Development of Nuclear gene-derived Molecular Phylogenetic Markers
Linked to Legume Genetic Maps
Unpublished (2004)

Contact: Douglas R. Cook
CAES Genomics Facility
University of California-Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Primer A: GGACTTTCTGCTGACCCAGAGACA
Primer B: ACCAAGGATGCTTTGTGGCATGGA
STS size: 166
PCR Profile:
Presoak: 95 degrees C for 15.00 minutes
Denaturaton: 94 degrees C for 20 seconds
Annealing: 55 degrees C for 20 seconds
Polymerization: 72 degrees C for 1.00 minutes
PCR Cycles: 35
Thermal cycler: MJ Research, Tetrad
Protocol:
Template: 5-30 ng
Primer: each 0.5 uM
dNTP: each 250 uM
Tag Polymerase: 0.05 units/uI


```

Total Vol: 10 ul

Buffer:
MgCl2 2.0 mM
KCl 50 mM
Tris-HCl 10 mM
pH 8.7.

FEATURES
    source
        Location/Qualifiers
            1..166
                /organism="Vigna radiata"
                /mol_type="genomic DNA"
                /cultivar="TC1966"
                /db_xref="taxon:157791"
                /sex="Hermaphrodite"
                /tissue_type="Leaf"
                /clone_lib="PCR fragment of the molecular marker,
                MsU63-Vr-TC1966"
                /note="PCR-amplified genomic fragment using cross-species
                gene-specific primer sets"
            1..166
                /gene="MsU63"
                /note="Light harvesting chlorophyll a/b-binding (CAB)
                protein"
            <1..>166
                /gene="MsU63"

    STS

    ORIGIN

Query Match 10.7%; Score 29.2; DB 11; Length 166;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 73; Conservative 0; Mismatches 73; Indels 0; Gaps 0

QY 119 TGGGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTT 178
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21 TGGGGTGTGTGTTCCCTGAACTGTTGGTCCGGAACGGTGTGAAGTTCGGGAGGCAGTGT 80
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 179 CTGGCAGGGCTGCAGGAAAGGCTGTACAGCTGCGCTGAGATTGTCAAGCGGCGGTCCCAG 238
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 81 GGTTCAGGCTGGGTCTCAGATATTCAGTAGGGTGGGCTTGTCTACTTGGGCAACCCAA 140
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 239 GCCTGCACCAGCTCACCAAGCTACGT 264
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 GCCTGATCCATGCCACAAAGCATCCT 166
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 2
AB126591/c
LOCUS          133 bp      DNA      linear      MAM 20-MAY-2004
DEFINITION    Canis familiaris DRD4 gene, intron 2, complete sequence, isolate:Q.
ACCESSION     AB126591
VERSION       AB126591.1  GI:47522547
KEYWORDS
SOURCE
ORGANISM      Canis familiaris (dog)
               Canis familiaris
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE     1
AUTHORS       Nara,H. and Inoue-Murayama,M.
TITLE         Novel Polymorphism of the canine dopamine receptor D4 gene intron
               II region
JOURNAL       Unpublished
REFERENCE     2  (bases 1 to 133)
AUTHORS       Nara,H. and Inoue-Murayama,M.
TITLE         Direct Submission
JOURNAL       Submitted (19-NOV-2003) Miho Inoue-Murayama, Gifu-University,
               Faculty of Agriculture; Yanagido 1-1, Gifu, Gifu 501-1193, Japan
               (E-mail:miho-i@cc.gifu-u.ac.jp, Tel:81-58-293-2874(ex.2874),
               Fax:81-58-293-2874)
FEATURES
             Location/Qualifiers
             1..133
                /organism="Canis familiaris"
                /mol_type="genomic DNA"
                /isolate="Q"
                /db_xref="taxon:9615"
source

```

```

gene
intron

1. .133
/gene="DRD4"
1. .133
/gene="DRD4"
/note="dopamine receptor D4"
/number=2

ORIGIN
Query Match          10.5%;   Score 28.8;   DB 4;   Length 133;
Best Local Similarity 54.8%;   Pred. No. 1.9e+03;
Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY      137 GGTGGAGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAA 196
        ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       116 GGGCGGCGGCCGCGGGCGGGCCCTGGCAGGGCGGGCGAGGGCGGGCGCTGGCGAG 57

QY      197 AGGCTGTACAGTCGCTGAGATTGTCAAGCGCGGGTCCCAGGC 240
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       56 GGGCGGTGCGGAGCGGGCCCTGGCGAGGGCGGGCGGGCCGACGC 13

RESULT 3
BV200943/c
LOCUS
DEFINITION sqm205467 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
ACCESSION BV200943
VERSION BV200943.1 GI:48169634
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 201)
         Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 201.

FEATURES             location/Qualifiers
source               1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
                     <1. .>201

STS
ORIGIN
Query Match          10.5%;   Score 28.8;   DB 11;   Length 201;
Best Local Similarity 60.0%;   Pred. No. 1.9e+03;
Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY      72 TACCCTTGAGATCGGGTCCGAGATGGCAGCAAATTCGCAACCTGCTGGGGTTGGCTCT 131
        ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       200 TAGGTTTCAGACGCACGACCTTAGGGCAGGAGAACCAACCATCCGCTTTTGTGCT 141

QY      132 GGGTCGGTTGGAGGGCGGCA 151
        ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       140 AGGCGGTGGGATGCCGTCA 121

RESULT 4
```

RESULT 4

HSA506605/c
LOCUS HSA506605 227 bp DNA linear PRI 07-APR-2003
DEFINITION Homo sapiens partial IGLV3 gene, out of frame rearrangement, isolate case 2, GCE 149.
ACCESSION AJ506605
VERSION AJ506605.1 GI:27650792
KEYWORDS IGLV3 gene; immunoglobulin lambda light chain; variable region.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Kurth,J., Hansmann,M.L., Rajewsky,K. and Kuppers,R. Epstein-Barr virus-infected B cells expanding in germinal centers of infectious mononucleosis patients do not participate in the germinal center reaction
Proc. Natl. Acad. Sci. U.S.A. 100 (8), 4730-4735 (2003)
22584347
12665622
PUBMED
REFERENCE 2 (bases 1 to 227)
AUTHORS Kurth,J.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2002) Kurth J., Department of Internal Medicine I, University of Cologne, Joseph-Stelzmannstr. 9, Cologne, 50931, GERMANY
FEATURES
source Location/Qualifiers
1..227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="case 2, GCE 149"
/db_xref="taxon:9606"
/cell_type="B-lymphocyte"
/tissue_type="tonsil"
/rearranged
/note="EBER-positive cell, located in interfollicular region"
1..227
/gene="IGLV3"
/pseudo
/pseudo
1..227
/gene="IGLV3"
/note="out of frame rearrangement of immunoglobulin kappa chain variable region"
/pseudo
/codon_start=1
/db_xref="PSEUDO:CAD53530.1"
1..227
/gene="IGLV3"
/pseudo
V_region
ORIGIN
Query Match 10.5%; Score 28.8; DB 9; Length 227;
Best Local Similarity 49.3%; Pred. No. 1.9e+03;
Matches 75; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 100 AGCAAAATTCGCAACTGCTGGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGG 159
Db 186 AGGATAGCCGGCCTCATTCGCCGCTTCGACCCCTGCTGATGGTCAGGGTGCGGTTCTC 127
QY 160 CATGTAGTGTCTCAGGTTCTGGCAGGGCTGCAGGAAGCTGTCAAGTCGCGCTGAGATT 219
Db 126 AGAGTTGGAGTCAGAGAAATCGGTCAGCGATCCCTGAGGGCCGGTCGGCATCGTCATGGAC 67
QY 220 GTCAAGCGGGCGGTCCAGGCCTGCACCAGCT 251
Db 66 GACCAGCACAGGGGCCTGGCCTGGCTTCTGCT 35
RESULT 5
CQ660987/c
LOCUS CQ660987 262 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 5913 from Patent WO2070737.
ACCESSION CQ660987

VERSION CQ660987.1 GI:42133182
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 5913 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1..262
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 10.4%; Score 28.4; DB 6; Length 262;
Best Local Similarity 58.1%; Pred. No. 2.5e+03;
Matches 50; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 63 ACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTGGG 122
Db 86 ACTTCTCTGTAGGCTTCAGACGCACGACCTTGAGGGCAGCAGAACCATCCGCTTTT 27
QY 123 GTTGGCTCTGGTCTGGTTGGAGGGCG 148
Db 26 TCTTGTCTAGGGCGGTGGGATGCCG 1
RESULT 6
MMSTS122/c
LOCUS MMSTS122 273 bp DNA linear STS 25-OCT-1995
DEFINITION M.musculus (D11Bhm122) sequence tagged site DNA.
ACCESSION Z36745
VERSION Z36745.1 GI:1041471
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 273)
AUTHORS Nehls,M., Luno,K., Schorpp,M., Pfeifer,D., Krause,S., Matysiak-Scholze,U., Dierbach,H. and Boehm,T.
TITLE YAC/Pl contigs defining the location of 56 microsatellite markers and several genes across a 3.4-cM interval on mouse chromosome 11
JOURNAL Mamm. Genome 6 (5), 321-331 (1995)
MEDLINE 95352957
PUBMED 7626882
REFERENCE 2 (bases 1 to 273)
AUTHORS Boehm,T.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 273)
AUTHORS Boehm,T.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1994) Thomas Boehm, Dept. of Medicine I, Molecular Medicine Group, Hugstetter Str. 55, Freiburg, 79106, Germany
FEATURES
source Location/Qualifiers
1..273
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
ORIGIN
Query Match 10.3%; Score 28.2; DB 11; Length 273;
Best Local Similarity 57.3%; Pred. No. 2.9e+03;
Matches 51; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 76 CTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTGGGTTGGCTCTGGGT 135

Db 200 CTTGAGAGGCAGCTGCAAGGGCGGCAGCTGTGGAAGGCAGCTGCCGGCTTGATCTCTAG 141

QY 136 CGGTTGGAGGGCGGCAGTGTCTGGGCATGT 164

Db 140 CCGGTATTATCTGACAGCTCTCGGCATGT 112

RESULT 7

MUSNR2C01

LOCUS MUSNR2C01 273 bp DNA linear ROD 23-JAN-1995

DEFINITION Mouse N-methyl-D-aspartate receptor subunit NR2C (NMDA2C) gene, exon 1.

ACCESSION L35014

VERSION L35014.1 GI:602738

KEYWORDS N-methyl-D-aspartate receptor subunit NR2C; NMDA receptor subunit NR2C.

SEGMENT 1 of 16

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 273)

AUTHORS Suchanek,B., Seeburg,P.H. and Sprengel,R.

TITLE Gene structure of the murine N-methyl D-aspartate receptor subunit NR2C

JOURNAL J. Biol. Chem. 270 (1), 41-44 (1995)

MEDLINE 95113860

PUBMED 7814402

COMMENT Original source text: Mus musculus (strain SV129) (tissue library: Stratagene Lambda FIX II) adult liver DNA.

FEATURES

source Location/Qualifiers

1..273

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="SV129"

/db_xref="taxon:10090"

/tissue_type="liver"

/dev_stage="adult"

/tissue_lib="Stratagene Lambda FIX II"

11..263

/gene="NMDA2C"

/note="5' untranslated exon; NR2C mouse gene; transcriptional start sites at basepairs 11, 12, 29"

/number=1

exon

ORIGIN

Query Match 10.3%; Score 28; DB 10; Length 273;

Best Local Similarity 56.5%; Pred. No. 3.3e+03;

Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 90 CCGAGATGGCAGCAAAATTGCAACCTGTGGGTTGGCTCTGGTCTGGTTGGAGGGCGG 149

Db 159 CCGAGCTCGCAGCCAGCGTGCAGCTCACGGGGCTGGCTCTGGACGAGCACCGCGCGG 218

QY 150 CAGTGCTCGGCATGTAGTGTCTCAGGTTCTG 181

Db 219 GAGACGGCGCGGGCTCTGGGCGTGGTGCCAG 250

RESULT 8

CQ679231/c

LOCUS CQ679231 154 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 24157 from Patent WO02070737.

ACCESSION CQ679231

VERSION CQ679231.1 GI:42191392

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1

AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.

TITLE Compositions and methods relating to osteoarthritis

Db 200 CTTGAGAGGCAGCTGCAAGGGCGGCAGCTGTGGAAGGCAGCTGCCGGCTTGATCTCTAG 141

QY 136 CGGTTGGAGGGCGGCAGTGTCTGGGCATGT 164

Db 140 CCGGTATTATCTGACAGCTCTCGGCATGT 112

RESULT 7

MUSNR2C01

LOCUS MUSNR2C01 273 bp DNA linear ROD 23-JAN-1995

DEFINITION Mouse N-methyl-D-aspartate receptor subunit NR2C (NMDA2C) gene, exon 1.

ACCESSION L35014

VERSION L35014.1 GI:602738

KEYWORDS N-methyl-D-aspartate receptor subunit NR2C; NMDA receptor subunit NR2C.

SEGMENT 1 of 16

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 273)

AUTHORS Suchanek,B., Seeburg,P.H. and Sprengel,R.

TITLE Gene structure of the murine N-methyl D-aspartate receptor subunit NR2C

JOURNAL J. Biol. Chem. 270 (1), 41-44 (1995)

MEDLINE 95113860

PUBMED 7814402

COMMENT Original source text: Mus musculus (strain SV129) (tissue library: Stratagene Lambda FIX II) adult liver DNA.

FEATURES

source Location/Qualifiers

1..273

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="SV129"

/db_xref="taxon:10090"

/tissue_type="liver"

/dev_stage="adult"

/tissue_lib="Stratagene Lambda FIX II"

11..263

/gene="NMDA2C"

/note="5' untranslated exon; NR2C mouse gene; transcriptional start sites at basepairs 11, 12, 29"

/number=1

exon

ORIGIN

Query Match 10.3%; Score 28; DB 10; Length 273;

Best Local Similarity 56.5%; Pred. No. 3.3e+03;

Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 90 CCGAGATGGCAGCAAAATTGCAACCTGTGGGTTGGCTCTGGTCTGGTTGGAGGGCGG 149

Db 159 CCGAGCTCGCAGCCAGCGTGCAGCTCACGGGGCTGGCTCTGGACGAGCACCGCGCGG 218

QY 150 CAGTGCTCGGCATGTAGTGTCTCAGGTTCTG 181

Db 219 GAGACGGCGCGGGCTCTGGGCGTGGTGCCAG 250

RESULT 8

CQ679231/c

LOCUS CQ679231 154 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 24157 from Patent WO02070737.

ACCESSION CQ679231

VERSION CQ679231.1 GI:42191392

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1

AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.

TITLE Compositions and methods relating to osteoarthritis

JOURNAL Patent: WO 02070737-A 24157 12-SEP-2002;

Chondrogene Inc. (CA)

FEATURES

source Location/Qualifiers

1..154

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 10.2%; Score 27.8; DB 6; Length 154;

Best Local Similarity 56.2%; Pred. No. 3.7e+03;

Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 63 ACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTGGG 122

Db 153 ACTTCTGTAGGCTTCANACGCCGACCTTNAGGGCAGCAGGACCCACCATCCGCTTTT 94

QY 123 GTTGGCTCTGGGTGCGTTGGAGGGCGGCA 151

Db 93 TCTTGTCTAGGGCGGTGGGATGCCGTCA 65

RESULT 9

CQ443803

LOCUS CQ443803 194 bp DNA linear PAT 30-JAN-2004

DEFINITION Sequence 9563 from Patent WO0192523.

ACCESSION CQ443803

VERSION CQ443803.1 GI:41416132

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1

AUTHORS Shimkets,R.A. and Leach,M.D.

TITLE Human polynucleotides and polypeptides encoded thereby

JOURNAL Patent: WO 0192523-A 9563 06-DEC-2001;

Curagen Corporation (US)

FEATURES

source Location/Qualifiers

1..194

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 10.2%; Score 27.8; DB 6; Length 194;

Best Local Similarity 59.5%; Pred. No. 3.7e+03;

Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 171 CTCAGGTTCTGGCAGGCTGCAGGAAGGCTGTCTGAGCTGCGCTGAGATTGTCAAGCGGCG 230

Db 11 CTCAGGTTCTGGGCGGTGGAGTAAGCAGGCAGATGAGCTGCCCTGAGATTGCTGGGAGGAG 70

QY 231 GGTCCCAGGCGCTGCACCAG 249

Db 71 GGAAGCGGGCAGACCCAG 89

RESULT 10

HSA230371/c

LOCUS HSA230371 240 bp DNA linear PRI 02-MAR-1999

DEFINITION Homo sapiens BF2N3-L3-C11 gene for immunoglobulin lambda chain variable region.

ACCESSION AJ230371

VERSION AJ230371.1 GI:4468513

KEYWORDS immunoglobulin; immunoglobulin lambda chain; immunoglobulin superfamily; joining region; variable region.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 240)

AUTHORS Farner,N.L.

National Institute of Agrobiological Resources
Kannondai 2-1-2
Tsukuba, Ibaraki
Japan
Phone: 0298-38-7441
Fax: 0298-38-7468
PROJECT = 'RGP'

FEATURES source
1. .160
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

ORIGIN

Query Match 9.9%; Score 27; DB 11; Length 160;
Best Local Similarity 56.0%; Pred. No. 6.3e+03;
Matches 51; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 130 CTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCTGGCAGGGCT 189
Db 126 CTCGGCTGTCGAGCGCGCGTGGTGATCTTGTGGCTTCTTGGTTGGCGTCGAA 67
QY 190 GCAGGAAAGCTGTGTCAGTCGCGCTGAGATTG 220
Db 66 GCAGCAGTCGATGTCCTCGATGACGAGGATG 36

RESULT 14
HS140C7F 241 bp DNA linear PRI 19-OCT-1995
LOCUS H.sapiens CpG island DNA genomic MseI fragment, clone 140C7,
DEFINITION forward read cp9140c7.ftla.
ACCESSION Z59291
VERSION Z59291.1 GI:1031204
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
REFERENCE 2 (bases 1 to 241)
AUTHORS Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biobelp@hgmp.mrc.ac.uk.

FEATURES source
1. .241
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="140C7"
/sex="male"
/tissue type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"

ORIGIN

Query Match 9.9%; Score 27; DB 9; Length 241;
Best Local Similarity 51.7%; Pred. No. 6.4e+03;
Matches 60; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 60 GCTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCT 119

Db 122 GGTCCACTTGTGGCGGAAGAGGGGCTCTGTGNCCTAGATGCCCTAGTTCA 181
QY 120 GGGTTGGCTCTGGTTCGGTTGGAGGGCGCAGTCTCGGCATGTAGTGTCTCAG 175
Db 182 GGCCCGGCTTGGGCGCTGGAGGGACCGCTTCCGTAGCTAGAGTTCTCCG 237

RESULT 15
BD237006 264 bp DNA linear PAT 17-JUL-2003
LOCUS Compounds for remedy and diagnosis of lung cancer and method for
DEFINITION using the same.
ACCESSION BD237006
VERSION BD237006.1 GI:33046776
KEYWORDS JP 2002516659-A/7.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 264)
AUTHORS Reed,S.G., Lodes,M.J., Frudakis,T.N. and Mohamath,R.
TITLE Compounds for remedy and diagnosis of lung cancer and method for
JOURNAL Patent: JP 2002516659-A 7 11-JUN-2002;
COMMENT CORIXA CORP
OS Homo sapiens (human)
PN JP 2002516659-A/7
PD 11-JUN-2002
PF 26-JAN-1999 JP 2000529432
PR 28-JAN-1998 US 09/015029,28-JAN-1998 US 09/015022 PR
18-MAR-1998 US 09/040828,18-MAR-1998 US 09/040831 PR
23-JUL-1998 US 09/122192,23-JUL-1998 US 09/122191 PR
22-DEC-1998 US 09/219245

PI STEVEN G REED,MICHAEL J LODES,TONY N FRUDAKIS,RAODOH MOHAMATH
PC C12N15/09,A61K35/14,A61K38/00,A61K39/00,A61K39/39,A61K39/395,
PC A61K39/395,
PC A61P11/00,A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,G01N33/577//
PC (C12N1/21,C12R1:19),C12N15/00,A61K37/02,C12N5/00 CC n =
A,T,C or G
FH Key Location/Qualifiers
FT misc feature (1)..(264).

FEATURES source
1. .264
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 9.9%; Score 27; DB 6; Length 264;
Best Local Similarity 54.3%; Pred. No. 6.4e+03;
Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 175 GGTCTGGCAGGGCTGCAGGAAAGGCTGTTCAGTGGCTGAGATTGTCAAGCGGGGTC 234
Db 12 GGATGGGGCACGGGAGCACAGATMGACTTTAACTGCCCCACGTTNTCMAGGAAAGGATT 71
QY 235 CCAGGCTTCACCCAGCTCACCAGCTACGTTTCC 268
Db 72 ACAGGCGTGAGCCACTGCGCCCGGCTCTTCTCC 105

Search completed: January 31, 2005, 19:44:40
Job time : 1826 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 18:58:47 ; Search time 1938 Seconds
(without alignments)
5133.151 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaaagctacgtttccttcag 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 7660530

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	225.4	82.6	272	4	BM699759	BM699759 UI-E-DW1-
2	206.2	75.5	268	1	AA022458	AA022458 ze70h07.r
3	138.4	50.7	194	2	AW436738	AW436738 77186 MAR
4	135	49.5	220	6	CB117372	CB117372 K-EST0162
5	117.4	43.0	251	7	W48348	W48348 mc86b02.r1
6	103	37.7	256	1	AI183131	AI183131 ub93g04.r
7	101.6	37.2	217	1	AA239107	AA239107 my36c09.r
8	97.6	35.8	246	7	W07888	W07888 mb49d03.r1
9	96.4	35.3	188	1	AA756051	AA756051 vull1e10.r
10	78.2	28.6	120	1	AI391194	AI391194 mb49d03.y
11	46.6	17.1	263	2	AW326399	AW326399 18938 MAR
12	33.8	12.4	247	4	BM068792	BM068792 ie85a09.y
13	33	12.1	230	1	AV435580	AV435580 AV435580
14	32.4	11.9	220	2	AW818811	AW818811 RC4-ST027
15	32.2	11.8	133	1	AV430042	AV430042 AV430042
16	32.2	11.8	181	1	AV432681	AV432681 AV432681
17	32.2	11.8	204	1	AV434181	AV434181 AV434181
18	32.2	11.8	206	1	AV435325	AV435325 AV435325
19	32.2	11.8	210	1	AV432576	AV432576 AV432576
20	32.2	11.8	227	1	AV433480	AV433480 AV433480
21	32.2	11.8	234	1	AV436270	AV436270 AV436270
22	31.6	11.6	231	2	BF478851	BF478851 WHE2014.D
23	31.4	11.5	240	2	BE774505	BE774505 PM0-UM001
24	31.4	11.5	250	2	BE771297	BE771297 QV1-FT008

25	31.4	11.5	258	6	CB885146	CB885146 Ma0060 Ha
C 26	31.4	11.5	262	4	BG953905	BG953905 RC4-CT069
C 27	31.2	11.4	268	2	BE937539	BE937539 RC4-ST027
28	31	11.4	247	2	BE717545	BE717545 RC4-HT077
C 29	31	11.4	260	7	CN363918	CN363918 170005999
C 30	30.8	11.3	233	1	AV435417	AV435417 AV435417
C 31	30.8	11.3	246	1	AV435673	AV435673 AV435673
C 32	30.8	11.3	249	1	AV432026	AV432026 AV432026
C 33	30.8	11.3	266	9	FR0044635	AL132127 Fugu rubr
C 34	30.6	11.2	135	1	AV430758	AV430758 AV430758
C 35	30.6	11.2	229	4	BJ162793	BJ162793 BJ162793
C 36	30.4	11.1	121	1	AV435623	AV435623 AV435623
C 37	30.4	11.1	188	1	AL045467	AL045467 DKF2p434H
38	30.4	11.1	223	7	CO264753	CO264753 EK140247.
C 39	30.4	11.1	224	1	AA301864	AA301864 EST15184
C 40	30.4	11.1	230	1	AV435742	AV435742 AV435742
41	30.4	11.1	236	4	BM740588	BM740588 K-EST0011
C 42	30.4	11.1	241	1	AV435395	AV435395 AV435395
C 43	30.2	11.1	232	1	AI508914	AI508914 vc46e05.y
C 44	30.2	11.1	251	7	R25661	R25661 Yh47e06.s1
C 45	30.2	11.1	252	7	CO698646	CO698646 DG32-142c

ALIGNMENTS

RESULT 1
BM699759
LOCUS BM699759 272 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DW1-agz-i-16-0-UI.r1 UI-E-DW1 Homo sapiens cDNA clone
UI-E-DW1-agz-i-16-0-UI 5', mRNA sequence.
ACCESSION BM699759
VERSION BM699759.1 GI:19013017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 272)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRE, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source
1..272
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW1-agz-i-16-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW1 is a normalized cDNA library containing the
following tissue(s): lens. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 82.6%; Score 225.4; DB 4; Length 272;
Best Local Similarity 99.1%; Pred. No. 1.5e-51;
Matches 226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
45 ATTGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 104

QY 61 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGCCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
105 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGCCAGCAAAATTCGCAACCTGCTG 164

QY 121 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTTTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
165 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTTTCTCAGGTTCT 224

QY 181 GCGAGGGCTGCAGGAAAGCTGTCTCAGCTGCGTGTAGATTGTCAAGCGG 228
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
225 GCGAGGGCTGCAGGAAAGCTGTCTCAGCTGCGTGTAGATTGTCAAGCGG 272

RESULT 2

AA022458
ze70h07.r1 Soares fetal_heart_NbHH19W Homo sapiens cDNA clone
IMAGE:364381 5', mRNA sequence.

AA022458
AA022458.1 GI:1486557
EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 945 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 200.
Location/Qualifiers
1. .268
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source

/db_xref="GDB:1281085"
/db_xref="taxon:9606"
/clone="IMAGE:364381"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal_heart_NbHH19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ORIGIN

Query Match 75.5%; Score 206.2; DB 1; Length 268;
Best Local Similarity 97.7%; Pred. No. 3.1e-46;
Matches 208; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
54 ATGGAGCACTACCGGAAAGCTGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG 113

QY 61 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGCCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 CTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGCCAGCAAAATTCGCAACCTGCTG 173

QY 121 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTTTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 GGGTTGGCTCTGGGTGGAGGGCGGAGTGTCTGGCATGTAGTTTCTCAGGTTCT 233

QY 181 GCGAGGGCTGCAGGAAAGCTGTCTCAGCTGCGCT 213
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 GCGAGGGCTNCAGGAAAGCTGTCTGCT 266

RESULT 3

AW436738

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCAGACG
Plate: 32 row: M column: 15
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source
1. .194
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC'2PIG"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Query Match 50.7%; Score 138.4; DB 2; Length 194;
Best Local Similarity 90.2%; Pred. No. 1.6e-27;
Matches 148; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG 60
|||
Db 30 ATGGAGCACTACCGGAAAGCTGGCTCTGTGGAATACTACCAGCACCTTCCCCAATGCCCCAG 89
QY 61 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 120
|||
Db 90 CTGCCTCTGATACCTTGAGATCGAGTCCGAGATGGCAGCAGATCCGCAACCTGCTG 149
QY 121 GGGTTGGCTCTGGTGGGTTGGAGGCGGCAGTGTCTGGCATGT 164
|||
Db 150 GGACTGGCGCTGGGGCGTTGGAGGTTGGCAGCGCCCGGCATGT 193

RESULT 4

CB117372
LOCUS K-EST0162559 L4SNU368 Homo sapiens CDNA clone L4SNU368-30-H04 5',
DEFINITION mRNA sequence.
ACCESSION CB117372
VERSION CB117372.1 GI:27943179
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 220)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

21c Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 30 row: H column: 04

High quality sequence stop: 220.

Location/Qualifiers

1. .220
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L4SNU368-30-H04"
/sex="M"
/tissue_type="Liver"
/cell_type="Polygonal"
/cell_line="SNU-368"
/lab_host="Top10F"
/clone_lib="L4SNU368"
/note="Organ: Liver; Vector: pcNS-D2; Site_1: EcoRI;

FEATURES

source

Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 49.5%; Score 135; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.4e-26;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG 60
|||
Db 86 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG 145
QY 61 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 120
|||
Db 146 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTG 205
QY 121 GGGTTGGCTCTGGGT 135
|||
Db 206 GGGTTGGCTCTGGGT 220

RESULT 5

W48348
LOCUS W48348
DEFINITION mc86b02.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:355371 5', mRNA sequence.

ACCESSION W48348

VERSION W48348.1 GI:13333399

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 251)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

UNPUBLISHED (1996)

CONTACT: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:227171

Seq primer: mob.REGA+ET

High quality sequence stop: 201.

Location/Qualifiers

1. .251

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:355371"

FEATURES

source

/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5,
TGTTACCAATCTGAAGTGGGAGCGCGCGGAATTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

ORIGIN

Query Match 43.0%; Score 117.4; DB 7; Length 251;
Best Local Similarity 82.1%; Pred. No. 1e-21;
Matches 147; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 74 ATGGAACAGTACCGG-AGGCCGGCTCTGTAGAGCTTCAGCCTCATCACCAATGCCCCAG 132

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 133 CTACCTCCTGACACTCTGGAATGCGAGTCCGAGATGGGAGCAAAATCCGAAACCTGCTA 192

QY 121 GGGTTGGCTCTGGTTCGGTGGAGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTTC 179
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 193 GGGCTGGCGTGGTTCGGTTCGGTGGAGGGGGAAGCACCAGGCACGTGGTGTCTCAGGCTC 251

RESULT 6

AI183131
LOCUS ub93q04.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
DEFINITION IMAGE:1396086 5', mRNA sequence.
AI183131
ACCESSION AI183131.1 GI:3733769
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 256)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:907802
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 252.
Location/Qualifiers
1. .256
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

FEATURES

source

/db_xref="taxon:10090"
/clone="IMAGE:1396086"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary_gland_NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5,
TGTTACCAATCTGAAGTGGGAGCGCGCGGAATTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN

Query Match 37.7%; Score 103; DB 1; Length 256;
Best Local Similarity 81.7%; Pred. No. 9.9e-18;
Matches 143; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84 ATGGAACAGTACCGG-AGGCCGGCTCTGTAGA-CTTCCAGCCTCATCACCAATGCCCCAG 141

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 142 CTACCTCCTGACACTCTGGAATGCGAGTCCGAGATGGGAGCAAAATCCGAAACCTGCTA 201

QY 121 GGGTTGGCTCTGGTTCGGTGGAGGCGGCAGTGTCTGGCATGTAGTGTCTCTCAG 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 202 GGGCTGGCGTGGTTCGGTTCGGTGGAGGGGGAAGCACCAGGCACGTGGTGTCTCTCAG 256

RESULT 7

AA239107
LOCUS my36c09.r1 Barstead mouse pooled organs MPLRB4 Mus musculus CDNA
DEFINITION clone IMAGE:697936 5', mRNA sequence.
AA239107
ACCESSION AA239107.1 GI:1863130
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 217)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:431496
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 169.
Location/Qualifiers
1. .217
/organism="Mus musculus"
/mol_type="mRNA"

FEATURES

source

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Query Match 17.1%; Score 46.6; DB 2; Length 263;
Best Local Similarity 92.5%; Pred. No. 0.038; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 4;
QY 1 ATGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCAAT 53
|||||
Db 211 ATGGAGCACTACCGGAAGCTGGCTCTGTGGACTACCAGCACCTTCCCAAT 263
|||||

RESULT 12

BM068792/c
LOCUS
DEFINITION ie85a09.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5673520 5', mRNA sequence.

ACCESSION

BM068792
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Other ESTs: ie85a09.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 221.

FEATURES

source

1. .247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5673520"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this
library."

ORIGIN

Query Match 12.4%; Score 33.8; DB 4; Length 247;
Best Local Similarity 60.2%; Pred. No. 1.3e+02;
Matches 56; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 36 CCCAGCGCCTTCCCCAATGCCCACTACTCCTGATACCCCTTGAGATGGGGTCCGAGA 95
|||||
Db 113 CCCGGGCTCCGGCCCTTCGCAGACCTGGCGCGGCGGTGCACATGGGGTCAAGGA 54
|||||
QY 96 TGGCAGCAAAATTGCAACCTGCTGGGGTTGGC 128
|||||
Db 53 AGGCAGCAAGATCCGTAACTGATGGCCTTCGC 21

RESULT 13

AV435580/c
LOCUS
DEFINITION AV435580 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM059e02_r 5', mRNA sequence.

ACCESSION

AV435580
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 230)
Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
DNA Res. 7, 223-227 (2000)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1. .230
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PM059e02_r"
/clone_lib="Porphyra yezoensis TU-1"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI"

ORIGIN

Query Match 12.1%; Score 33; DB 1; Length 230;
Best Local Similarity 54.5%; Pred. No. 2.1e+02;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 95 ATGGCAGCAAAATTCGCAACCTGCTGGGGTTGGCTCTGGGTCGTTGGAGGGCGGAGTG 154
|||||
Db 121 ATGACGGGACCTTGGGGCGGGCGGGGGCTAGGGACGGGGAGCCCGCGGGCTGG 62
|||||
QY 155 CTCGGCATGTAGTGTCTCAGGTTCTGCGAGGGGCTGCAGGAAAGGCTGTCAGCTGCGCTG 214
|||||
Db 61 CTGGGGATGGCGGGGGCTGGGCAGGTGGGTGGCGGGTGGATAGGGGGCGGAGGCTC 2
|||||
QY 215 A 215
|
Db 1 A 1

RESULT 14

AW818811/c

LOCUS AW818811 220 bp mRNA linear EST 17-MAY-2000
DEFINITION RC4-ST0279-010400-019-g04 ST0279 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW818811
VERSION AW818811.1 GI:79111909
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 220)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC4-ST0279-010400-019-g04&t3=2000-04-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 36
High quality sequence stop: 114.

FEATURES
source
1..220
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0279"
/note="Organ: stomach; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 11.9%; Score 32.4; DB 2; Length 220;
Best Local Similarity 60.0%; Pred. No. 3.1e+02;
Matches 54; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 63 ACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTGGG 122
||| ||||| ||| ||||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 190 ACTTTCGTGATAGGCTTCAGACGCACGACCTTGTGGGCAGCAGGAACCACTCCGCTTTT 131
QY 123 GTTGGCTCTGGTCGGTTGGAGGCGGCAG 152
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TCTTGTCTAGGCGGTGGGATGCCGTCAG 101

RESULT 15
AV430042/c
LOCUS AV430042 133 bp mRNA linear EST 23-AUG-2000
DEFINITION AV430042 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PL012e03_r 5', mRNA sequence.
ACCESSION AV430042
VERSION AV430042.1 GI:8585267
KEYWORDS EST.

SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
REFERENCE 1 (bases 1 to 133)
AUTHORS Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
PUBMED 10907854
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..133
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PL012e03_r"
/clone_lib="Porphyra yezoensis TU-1"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 11.8%; Score 32.2; DB 1; Length 133;
Best Local Similarity 54.7%; Pred. No. 3.3e+02;
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 95 ATGGCAGCAAAATTCGCAACCTGCTGGGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTG 154
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 ATGGCGCGACCCCTGGGGGCGGCGGGGGCTAGGACGGGAGCCCGCGGGGCTGG 60
QY 155 CTCGGCATGTAGTGTTCACGGTTCCTGGCAGGGCTGCAGAAAGGCTGTCAGCTGCG 211
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 CTGGGATGGCGGGGCTGGGCAGGTGGGTGGCGGGTGGATAGGGGGGCAGGTGAG 3

Search completed: January 31, 2005, 20:17:10
Job time : 1948 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 18:07:55 ; Search time 317 Seconds
(without alignments)
4520.795 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaagctacgttcttcag 273

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 4923180

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	29.8	10.9	192	6	ABQ55864	Abq55864 Human ova
C 2	28.6	10.5	265	12	ACH82956	Ach82956 Human gen
3	27.8	10.2	194	6	ABN20543	Abn20543 Human ORF
4	27.6	10.1	122	12	ACH85301	Ach85301 Human gen
C 5	27.2	10.0	133	12	ADQ91957	Adq91957 Dog dopam
C 6	27.2	10.0	235	10	ADF31067	Adf31067 Soil meta
7	27	9.9	264	3	AAC79052	Aac79052 Human lun
8	27	9.9	264	4	AAD23127	Aad23127 Human lun
9	27	9.9	264	10	ADD67064	Add67064 Human lun
10	27	9.9	264	10	ADE87569	Ade87569 Human lun
11	26.4	9.7	167	4	AAI26274	Aai26274 Probe #16
12	26.4	9.7	167	4	ABA73983	Aba73983 Human foe
13	26.4	9.7	167	4	AAI54432	Aai54432 Probe #23
14	26.4	9.7	167	4	ABA39062	Aba39062 Probe #17
15	26.4	9.7	167	4	AAK48603	Aak48603 Human bon
16	26.4	9.7	167	4	AAK22436	Aak22436 Human bra
17	26.4	9.7	167	4	ABS48287	Abs48287 Human liv
18	26.4	9.7	167	6	ABS22360	Abs22360 Human gen
C 19	26.4	9.7	198	4	AAI27005	Aai27005 Probe #16
C 20	26.4	9.7	198	4	ABA75273	Aba75273 Human foe
C 21	26.4	9.7	198	4	AAI55836	Aai55836 Probe #24

C 22	26.4	9.7	198	4	ABA39929	AbA39929 Probe #18
C 23	26.4	9.7	198	4	AAK49904	Aak49904 Human bon
C 24	26.4	9.7	198	4	AAK23818	Aak23818 Human bra
C 25	26.4	9.7	198	4	ABS49548	Abs49548 Human liv
C 26	26.4	9.7	198	6	ABS23400	Abs23400 Human gen
C 27	26.2	9.6	216	3	AAF10329	Aaf10329 Fusarium
C 28	26	9.5	198	4	ABA41025	AbA41025 Probe #19
C 29	26	9.5	249	8	ABT14594	Abt14594 Pseudomon
C 30	25.8	9.5	165	2	AAx51903	Aax51903 Human sec
31	25.8	9.5	169	6	ABL80211	AbL80211 Human ova
C 32	25.8	9.5	187	10	ACD94768	Acd94768 Human col
33	25.6	9.4	87	6	ADH31736	Adh31736 Yeast sma
C 34	25.6	9.4	212	10	ABx83988	Abx83988 Corn ear-
C 35	25.6	9.4	213	12	ACH86002	Ach86002 Human gen
C 36	25.6	9.4	261	11	ACH95344	Ach95344 Klebsiell
37	25.4	9.3	194	3	AAC18105	Aac18105 Human sec
38	25.4	9.3	235	10	ACA56063	Aca56063 Human sig
39	25.4	9.3	235	12	ADI55859	Adi55859 Human pol
C 40	25.4	9.3	262	9	ACH44670	Ach44670 Human foe
41	25.4	9.3	264	2	AAZ07150	Aaz07150 Human lun
C 42	25.2	9.2	109	12	ACH86303	Ach86303 Human gen
C 43	25.2	9.2	138	2	AAx11982	Aax11982 Human bia
C 44	25.2	9.2	219	8	ABx76158	Abx76158 Lung canc
C 45	25.2	9.2	219	11	ADP64955	Adp64955 Human sma

ALIGNMENTS

RESULT 1
ABQ55864/c
ID ABQ55864 standard; cDNA; 192 BP.

XX ABQ55864;

XX 22-AUG-2002 (first entry)

DE Human ovarian antigen HPAMQ76 cDNA, SEQ ID NO:1744.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.

XX Homo sapiens.

PN WO200200677-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001WO-US018569.

PR 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

DR P-PSDB; ABP42787.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases.

PS Claim 1; SEQ ID NO 1744; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 192 BP; 41 A; 57 C; 58 G; 36 T; 0 U; 0 Other;

Query Match 10.9%; Score 29.8; DB 6; Length 192;
Best Local Similarity 58.4%; Pred. No. 85;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 63 ACCTCCTGATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTGGG 122

Db 128 ACTTCTTGTAGGCTTCAGACGCACGACCTTGAGGGCAGCAGGAACCATCCGCTTTT 69

QY 123 GTTGGCTCTGGGTCGGTTGGAGGGCGGCA 151

Db 68 TCTTGCTGAGGCGGTGGGATGCCGTCA 40

RESULT 2
ACH82956/c
ID ACH82956 standard; DNA; 265 BP.

XX ACH82956;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #16151.

DE Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

XX (RANK/) RANK D R.

XX (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.

XX Claim 1; SEQ ID NO 16151; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX SQ Sequence 265 BP; 51 A; 97 C; 60 G; 57 T; 0 U; 0 Other;

Query Match 10.5%; Score 28.6; DB 12; Length 265;
Best Local Similarity 49.0%; Pred. No. 2.1e+02;
Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 2 TGGAGCACTACCGGAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAGC 61

Db 252 TCGAGATAGAGCTTCCACAGGGCTCTCTTCATGTCTCTCAGGCTGTAGATGAGC 193

QY 62 TACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTGG 121

Db 192 GGATTCAGGGTGGGTATGAGGATGGAGTAGAGCAGCAGCCAGCCCTTCCAGCGGG 133

QY 122 GGTGGCTCTGGGTGCGTTGGAGGGCGGCGAGTGCT 156

Db 132 GAGTAGCGGGTGGTGGTCCGATGTAGTGGAGCT 98

RESULT 3
ABN20543
ID ABN20543 standard; cDNA; 194 BP.
XX
AC ABN20543;
XX
DT 24-JUN-2002 (first entry)

XX	Human ORFX polynucleotide sequence SEQ ID NO:9563.	
DE		
XX		
KW	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;	
KW	hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;	
KW	degenerative disorder; osteoarthritis; neurodegenerative disorder;	
KW	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;	
KW	hypertension; hypothyroidism; cholesterol ester storage disease;	
KW	immune deficiency; immune disorder; infectious disease;	
KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;	
KW	myasthenia gravis; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200192523-A2.	
XX		
PD	06-DEC-2001.	
XX		
PF	29-MAY-2001; 2001WO-US010836.	
XX		
PR	30-MAY-2000; 2000US-0206132P.	
PR	29-AUG-2000; 2000US-0228716P.	
XX		
PA	(CURA-) CURAGEN CORP.	
XX		
PI	Shimkets RA, Leach MD;	
XX		
DR	WPI; 2002-106308/14.	
DR	P-PSDB; ABP04791.	
XX		
PT	Novel human polypeptides and polynucleotides useful for diagnosing,	
PT	preventing and treating cardiovascular disease, neurodegenerative,	
PT	hyperproliferative disorders and autoimmune disorders.	
XX		
PS	Disclosure; SEQ ID NO 9563; 1037pp; English.	
XX		
CC	The present invention describes substantially purified human proteins	
CC	(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1	
CC	in the specification). ABN15762 to ABN27252 encode the human ORFX	
CC	proteins given in ABP00010 to ABP11500. ORFX proteins are useful for	
CC	treating or preventing a pathology associated with an ORFX-associated	
CC	disorder in humans, and in the manufacture of a medicament for treating a	
CC	syndrome associated with ORFX-associated disorder. ORFX polynucleotide	
CC	sequences can be used in gene therapy. ORFX sequences can be used in the	
CC	treatment of cancer, hyperproliferative disorders, cirrhosis of liver,	
CC	psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,	
CC	osteoarthritis, neurodegenerative disorders, disorders related to organ	
CC	transplantation, cardiovascular diseases, diabetes mellitus, systemic	
CC	lupus erythematosus, hypertension, hypothyroidism, cholesterol ester	
CC	storage disease, various immune deficiencies and disorders, infectious	
CC	diseases, autoimmune disorders such as multiple sclerosis, rheumatoid	
CC	arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host	
CC	disease and autoimmune inflammatory eye disease. ORFX proteins are also	
CC	useful for treating burns, incisions, ulcers, for treating osteoporosis,	
CC	bone degenerative disorders, or periodontal disease, and for gut	
CC	protection or regeneration and treatment of lung or liver fibrosis,	
CC	reperfusion injury in various tissues and conditions resulting from	
CC	systemic cytokine damage. N.B. The sequence data for this patent did not	
CC	form part of the printed specification, but was obtained in electronic	
CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 194 BP; 41 A; 48 C; 74 G; 30 T; 0 U; 1 Other;	
	Query Match 10.2%; Score 27.8; DB 6; Length 194;	
	Best Local Similarity 59.5%; Pred. No. 3.3e+02;	
	Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;	
OY	171 CTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGTGCAGCTGCGCTGAGATTGTCAAGCGCG 230	
Db	11 CTCAGGGTCGGGGGTGGAGTAAGCAGGCAGATGAGCTGCCCTGAGATTGCTGGGAGGAG 70	
OY	231 GGTCCAGGCCTGCACGAG 249	

Db	71 GGAAGGGCGGCAGACCCAG 89	
	RESULT 4	
ACH85301		
ID	ACH85301 standard; DNA; 122 BP.	
XX		
AC	ACH85301;	
XX		
DT	29-JUL-2004 (first entry)	
XX		
DE	Human genome derived single exon probe #18496.	
XX		
KW	Human; probe; ss; gene expression; single exon probe; microarray;	
KW	alternative splicing event; genomic alteration.	
XX		
OS	Homo sapiens.	
XX		
PN	US2003194704-A1.	
XX		
PD	16-OCT-2003.	
XX		
PF	03-APR-2002; 2002US-00029386.	
XX		
PR	03-APR-2002; 2002US-00029386.	
XX		
PA	(PENN/) PENN S G.	
PA	(RANK/) RANK D R.	
PA	(HANZ/) HANZEL D K.	
XX		
PI	Penn SG, Rank DR, Hanzel DK;	
XX		
DR	WPI; 2004-119264/12.	
XX		
PT	New human genome-derived single exon nucleic acid probes useful for human	
PT	gene expression analysis, for identifying or characterizing alternative	
PT	splicing events, for assessing genomic alterations or as tools for	
PT	surveying tissues.	
XX		
PS	Claim 1; SEQ ID NO 18496; 80pp; English.	
XX		
CC	The invention relates to a nucleic acid probe for measuring human gene	
CC	expressions, comprising any of the 27,400 fully defined nucleotide	
CC	sequences in the specification, or their complements or fragments, and	
CC	encoding at least 8 amino acids of any of the 6888 amino acid sequences	
CC	fully defined in the specification. The probe is a single exon probe that	
CC	hybridises under high stringency conditions to a nucleic acid molecule	
CC	expressed in human cells or tissues. Also included are a spatially-	
CC	addressable set of single exon nucleic acid probes for measuring human	
CC	gene expression (comprising a plurality of single exon nucleic acid	
CC	probes cited above, where each of the plurality of probes is separately	
CC	and addressably isolatable or amplifiable from the plurality), a single	
CC	exon microarray for measuring human gene expression, a method of	
CC	measuring human gene expression, a vector comprising the single exon	
CC	probe cited above, an ORF-encoded peptide comprising at least 8	
CC	contiguous amino acids of any of the above-mentioned amino acid	
CC	sequences (optionally with conservative amino acid substitutions), an	
CC	isolated antibody that binds specifically to a peptide cited above,	
CC	methods of selling and/or licensing single exon probes or microarrays to	
CC	a customer desiring to measure gene expression, a method of providing	
CC	human gene expression data by subscription, and a computer-readable	
CC	storage medium which contains a database having a plurality of records	
CC	(each record including data on the expression of a single exon probe	
CC	cited above. The probe, methods and apparatus are useful in gene	
CC	expression analysis. The probes may be used as tools for surveying	
CC	tissues to detect the presence of expressed messages that contain their	
CC	specific exon, or in constructing genome-derived single exon microarrays.	
CC	In addition, the probes are used in identifying and characterising	
CC	alternative splicing events, in detecting and characterising gross	
CC	alterations in the genomic locus that includes their exon, in assessing	
CC	smaller genomic alterations, in priming the synthesis of nucleic acids,	
CC	or in expressing the ORF-encoded peptide. The present sequence is a human	
CC	single exon probe of the invention. Note: The sequence data for this	

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 122 BP; 17 A; 31 C; 39 G; 35 T; 0 U; 0 Other;
Query Match 10.1%; Score 27.6; DB 12; Length 122;
Best Local Similarity 58.5%; Pred. No. 3.3e+02;
Matches 48; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 128 CTCTGGTTCGGTTGGAGGCGGCAGTCTCGGCATGTAGTCTCAGGTTCTGCGCAGGG 187
Db 1 CTAGGATTCAGGTCGAATTCGCGCATGATCCGCTGCTGCAAGCCCAGGTTCTTACAGGA 60
QY 188 CTGCAGGAAGGCTGTCAGCTG 209
Db 61 ACGCTGACCGGGCTGCTGCTG 82
RESULT 5
ADQ91957/c
ID ADQ91957 standard; DNA; 133 BP.
XX
AC ADQ91957;
XX
DT 23-SEP-2004 (first entry)
XX
DE Dog dopamine receptor D4, DRD4, intron 2 long form.
XX
KW Dog; DRD4; intron 2; dopamine receptor D4; dog candidate; seeing-eye dog;
KW disaster rescue dog; assisting dog; narcotics detection dog; police dog;
KW animal herding dog; hunting dog; watchdog; housedog; ds.
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT allele 50..67
FT /*tag= a
FT /note= "This region is deleted in the short form"
XX
PN US2004121368-A1.
XX
PD 24-JUN-2004.
XX
PF 15-AUG-2003; 2003US-00641428.
XX
PR 24-DEC-2002; 2002JP-00373006.
XX
PA (UYGI-) UNIV GIFU.
XX
PI Murayama M, Ito S;
XX
DR WPI; 2004-479668/45.
XX
PT New polynucleotides and polypeptides, useful for screening useful dog
PT candidates having genetic aptitudes beneficial to people, or for
PT screening probes or identifying probes used in southern hybridization or
PT in situ hybridization.
XX
PS Claim 14; SEQ ID NO 8; 13pp; English.
XX
CC The invention relates to a polynucleotide derived from dog dopamine
CC receptor D4 (DRD4) comprising the short and long forms of exon 1
CC (ADQ91952 and ADQ91954) or the deleted portion from the long form
CC (ADQ91950), and the short and long forms of intron 2 (ADQ91957 and
CC ADQ91958 or the deleted portion of the long form (ADQ91956). Also
CC included are a transformant in which a polynucleotide is expressively
CC introduced that contains a nucleotide sequence selected from ADQ91950,
CC ADQ91952, ADQ91954, ADQ91956, ADQ91957 and ADQ91958, a polypeptide
CC containing an amino acid sequence encoded by the above nucleic acids
CC (ADQ91953 or ADQ91955, screening for useful dog candidates, a kit for
CC screening for useful dog candidates comprising a pair of PCR primers that
CC bonds to the vicinity of both terminals of an allele in exon 1 or intron

CC 2 of a dog dopamine receptor D4 gene, breeding useful dog candidates that
CC uses a dog in which the genotype of the allele in exon 1 or intron 2 of a
CC dog dopamine receptor D4 gene is homozygotic for at least one of the
CC parent dogs and a useful dog candidate that is selected based on the
CC genotype of the allele in exon 1 or intron 2 of a dog dopamine receptor
CC D4 gene. The service dogs are seeing-eye dogs, disaster rescue dogs,
CC assisting dogs, narcotics detection dogs, police dogs, dogs for the
CC hearing impaired, animal herding dogs, or hunting dogs. The dogs with
CC genetic aptitudes beneficial to people are watchdogs or housedogs. Dogs
CC having the short form of exon 1 are suited to being seeing eye dogs, dogs
CC having the short form of intron 2 display mild-mannered behavioural
CC traits and dogs having the long form of intron 2 display aggressive
CC behavioural traits. The polynucleotides and polypeptides are useful for
CC screening useful dog candidates having genetic aptitudes beneficial to
CC people. The polynucleotides can be used as screening probes or
CC identifying probes used in southern hybridisation, northern
CC hybridization, or in situ hybridisation. The present sequence is the Dog
CC dopamine receptor D4, DRD4, intron 2 long form.
XX
SQ Sequence 133 BP; 7 A; 83 C; 35 G; 8 T; 0 U; 0 Other;

Query Match 10.0%; Score 27.2; DB 12; Length 133;
Best Local Similarity 53.8%; Pred. No. 4.4e+02;
Matches 56; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 137 GGTTGGAGGCGGCAGTCTCGGCATGTAGTCTCAGGTTCTGCGCAGGGCTGCAGGAA 196
Db 116 GGGCGGCGGCGCGCGGGGCGGCTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 57
QY 197 AGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGGGGGTCCAGGC 240
Db 56 GGGCGGCTCGCAGGCGGGGGCGCTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 13

RESULT 6
ADF31067/c
ID ADF31067 standard; DNA; 235 BP.
XX
AC ADF31067;
XX
DT 12-FEB-2004 (first entry)
XX
DE Soil metagenome glycosyl hydrolase homologue DNA SEQ ID NO 287.
XX
KW glycosyl hydrolase; amylase; enzyme; washing; cleaning; textile;
KW raw material; cotton; starch liquefaction; ethanol production;
KW oligosaccharide preparation; cyclodextrin; hydrolysis; food; animal feed;
KW antistaling; additive; bread; adhesive; soil metagenome; ds.
OS Synechocystis sp.
XX
PN WO2003054177-A2.
XX
PD 03-JUL-2003.
XX
PF 13-DEC-2002; 2002WO-EP014210.
XX
PR 21-DEC-2001; 2001DE-01063748.
XX
PA (HENK) HENKEL KGAA.
XX
PI Breves R, Maurer K, Eck J, Lorenz P, Zinke H;
XX
DR WPI; 2003-559145/52.
XX
PT New glycosyl hydrolase, useful for degrading starch, for e.g. in washing
PT compositions, and in preparation of foods and animal feeds, e.g. as
PT antistaling additives for bread.
XX
PS Claim 34; SEQ ID NO 287; 275pp; German.
XX
CC This invention describes novel glycosyl hydrolases and their derivatives
CC which have amylase activity and are used (i) for fusing proteins

CC together, especially for developing new enzymes; (ii) in washing and
CC cleaning compositions for textiles or hard surfaces; (iii) for treating
CC raw materials and intermediates in preparation of textiles, especially
CC desizing of cotton; (iv) for starch liquefaction, especially in ethanol
CC production; (v) for preparation of linear and/or short-chain
CC oligosaccharides; (vi) for hydrolysis of cyclodextrins; (vii) for
CC releasing low molecular weight compounds from polysaccharide complexes or
CC cyclodextrins; (viii) in preparation of foods and animal feeds, e.g. as
CC antistaling additives for bread; (ix) for destroying starch-based
CC adhesives and (x) as temporary adhesives. Nucleic acid encoding the novel
CC glycosyl hydrolases and their derivatives can be fused to other nucleic
CC acid, especially for development of new enzymes. This sequence encodes a
CC soil metagenome glycosyl hydrolase homologue.

XX SQ Sequence 235 BP; 43 A; 88 C; 72 G; 32 T; 0 U; 0 Other;

Query Match 10.0%; Score 27.2; DB 10; Length 235;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 68; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 115 CTGCTGGGTTGGCTCTGGTTCGGTTCGGTGGCGGCGCAGTGCTCGGCATGTAGTGTCTCA 174
DB 205 CTCCTTGGCCCCAGGTGACGACCGAGTCGACGAGCTTGCCCATCATGGCGTTCCTCGGT 146
QY 175 GGTTCCTGGCAGGGCTGCAGGAAGGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGCGGGTTC 234
DB 145 GCGGTGTTCGAGCAGCAGGTGCTGTTGGCCACCGAGCCGTGCGCGAGCGCGTGGTA 86
QY 235 CCAGGCTCGCACCAGC 250
DB 85 GTAGCGGGCACGATC 70

RESULT 7

AAC79052
ID AAC79052 standard; cDNA; 264 BP.
XX
AC AAC79052;
XX
DT 05-FEB-2001 (first entry)
XX Human lung tumour-specific cDNA #7.
DE
XX Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
KW
XX Homo sapiens.
OS
XX WO200060077-A2.
PN
XX 12-OCT-2000.
PD
XX 30-MAR-2000; 2000WO-US008560.
PF
XX 02-APR-1999; 99US-00285323.
PR
XX 09-AUG-1999; 99US-00370838.
PR
XX 30-DEC-1999; 99US-00476235.
PR
XX 03-MAR-2000; 2000US-00518809.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX WPI; 2000-638466/61.
DR
XX Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT monitoring or treating cancer, especially lung cancer.
PT
XX Example 1; Page 86; 243pp; English.
PS
XX The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed. The
CC polypeptides are useful for inhibiting the development of cancer,

CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer
XX
SQ Sequence 264 BP; 52 A; 69 C; 63 G; 74 T; 0 U; 6 Other;

Query Match 9.9%; Score 27; DB 3; Length 264;
Best Local Similarity 54.3%; Pred. No. 6.1e+02;
Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 175 GGTTCCTGGCAGGGCTGCAGGAAGGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGCGGGTTC 234
DB 12 GGATGGGGCACGGGAGCACAGATMGACTTTAACTGCCCCCAGCTTNTCMAGGAAAGGATT 71
QY 235 CCAGGCTGCACCAGCTCACCAAGCTACGTTTCC 268
DB 72 ACAGGCGTGAGCCACTGCGCCCGGCTCTCTCC 105

RESULT 8

AAD23127
ID AAD23127 standard; cDNA; 264 BP.
XX
AC AAD23127;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific L164C1.cons cDNA.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; L164C1.cons;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US009991.
XX
PR 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00588937.
PR 18-AUG-2000; 2000US-00640878.
PR 22-SEP-2000; 2000US-0234517P.
PR 01-NOV-2000; 2000US-00704512.
PR 14-DEC-2000; 2000US-00738973.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR WPI; 2001-639201/73.
XX
PT New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer.
XX
PS Example 1; Page 143-144; 378pp; English.
XX
CC The invention relates to isolated lung tumour-specific proteins and their
CC corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is human
CC lung tumour-specific cDNA
XX
SQ Sequence 264 BP; 52 A; 69 C; 63 G; 74 T; 0 U; 6 Other;


```
RESULT 13
AAI54432
ID AAI54432 standard; DNA; 167 BP.
XX
AC AAI54432;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #23118 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 23118; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;

Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 114 CCTGCTGGGGTTGGCTCGGTCGGTTGGAGGCGGCAGTGTCTCGGCATGTAGTTCTC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 CCTGCTGGGGCTGGGGCTGGGGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTCTC 84

RESULT 14
ABA39062
ID ABA39062 standard; DNA; 167 BP.
XX
AC ABA39062;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #17528 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
```

```
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 17528; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;

Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 114 CCTGCTGGGGTTGGCTCGGTCGGTTGGAGGCGGCAGTGTCTCGGCATGTAGTTCTC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 CCTGCTGGGGCTGGGGCTGGGGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTCTC 84

RESULT 15
AAK48603
ID AAK48603 standard; DNA; 167 BP.
XX
AC AAK48603;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 23160.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PR 30-JAN-2001; 2001WO-US000668.
XX
PF 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
```

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 23160; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 167 BP; 19 A; 53 C; 52 G; 43 T; 0 U; 0 Other;

Query Match 9.7%; Score 26.4; DB 4; Length 167;
Best Local Similarity 65.0%; Pred. No. 8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 114 CCTGCTGGGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTC 173
Db ||||||| ||| ||||| | | ||||| | ||||| | ||||| | |||||
25 CCTGCTGGGGCTGGGGCTGGGGCTGGCGCTGGCTGTCTCATTTGTGCTGGCTGGTCTCTC 84

Search completed: January 31, 2005, 19:14:11
Job time : 320 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:44:48 ; Search time 337 Seconds
(without alignments)
4654.681 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaagctacgtttccttcag 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues 3938240
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.8	10.9	192	16	US-10-264-049-1744
C 2	28.6	10.5	201	18	US-10-719-993-27866
C 3	28.6	10.5	265	15	US-10-029-386-16151
C 4	28.4	10.4	201	17	US-10-741-601-25556
C 5	28.4	10.4	262	16	US-10-242-535A-5913
C 6	28.4	10.4	262	16	US-10-085-783A-5913
C 7	28	10.3	177	18	US-10-425-115-43846
C 8	28	10.3	227	17	US-10-767-701-21806
C 9	27.8	10.2	154	16	US-10-242-535A-24157
C 10	27.8	10.2	154	16	US-10-085-783A-24157
C 11	27.6	10.1	122	15	US-10-029-386-18496
C 12	27.6	10.1	201	17	US-10-741-601-19299
C 13	27.6	10.1	201	17	US-10-741-601-23785
C 14	27.2	10.0	133	17	US-10-641-428-8
C 15	27.2	10.0	235	18	US-10-872-874-287
C 16	27	9.9	201	17	US-10-741-601-4903
C 17	27	9.9	201	17	US-10-741-601-4912
C 18	27	9.9	201	17	US-10-741-601-25546
C 19	27	9.9	264	9	US-09-738-973-7
C 20	27	9.9	264	9	US-09-854-133-7
C 21	27	9.9	264	15	US-10-144-649A-7
C 22	26.8	9.8	201	17	US-10-741-601-25571
C 23	26.8	9.8	201	17	US-10-741-601-25581
C 24	26.4	9.7	167	9	US-09-864-761-24382
C 25	26.4	9.7	198	9	US-09-864-761-25249
C 26	26.4	9.7	231	17	US-10-437-963-29629
C 27	26.2	9.6	201	17	US-10-741-601-20919
C 28	26.2	9.6	201	18	US-10-719-993-21749
C 29	26.2	9.6	201	18	US-10-719-993-21750
C 30	26.2	9.6	201	18	US-10-719-993-34280
C 31	26.2	9.6	201	18	US-10-719-993-34282
C 32	26.2	9.6	201	18	US-10-719-993-53459
C 33	26.2	9.6	201	18	US-10-719-993-53460
C 34	26.2	9.6	216	18	US-10-653-047-2852
C 35	26	9.5	198	9	US-09-864-761-26345
C 36	26	9.5	198	15	US-10-054-435-30
C 37	26	9.5	201	18	US-10-719-993-16143
C 38	26	9.5	249	15	US-10-127-032-36
C 39	26	9.5	266	18	US-10-425-115-101750
C 40	25.8	9.5	169	9	US-09-867-701-3189
C 41	25.8	9.5	201	18	US-10-719-993-34124
C 42	25.8	9.5	243	18	US-10-425-115-24398
C 43	25.8	9.5	261	18	US-10-425-115-79105
C 44	25.6	9.4	87	14	US-10-083-357-194
C 45	25.6	9.4	201	17	US-10-741-601-19290

C 13	27.6	10.1	201	17	US-10-741-601-23785	Sequence 23785, A
C 14	27.2	10.0	133	17	US-10-641-428-8	Sequence 8, Appli
C 15	27.2	10.0	235	18	US-10-872-874-287	Sequence 287, App
C 16	27	9.9	201	17	US-10-741-601-4903	Sequence 4903, Ap
C 17	27	9.9	201	17	US-10-741-601-4912	Sequence 4912, Ap
C 18	27	9.9	201	17	US-10-741-601-25546	Sequence 25546, A
C 19	27	9.9	264	9	US-09-738-973-7	Sequence 7, Appli
C 20	27	9.9	264	9	US-09-854-133-7	Sequence 7, Appli
C 21	27	9.9	264	15	US-10-144-649A-7	Sequence 7, Appli
C 22	26.8	9.8	201	17	US-10-741-601-25571	Sequence 25571, A
C 23	26.8	9.8	201	17	US-10-741-601-25581	Sequence 25581, A
C 24	26.4	9.7	167	9	US-09-864-761-24382	Sequence 24382, A
C 25	26.4	9.7	198	9	US-09-864-761-25249	Sequence 25249, A
C 26	26.4	9.7	231	17	US-10-437-963-29629	Sequence 29629, A
C 27	26.2	9.6	201	17	US-10-741-601-20919	Sequence 20919, A
C 28	26.2	9.6	201	18	US-10-719-993-21749	Sequence 21749, A
C 29	26.2	9.6	201	18	US-10-719-993-21750	Sequence 21750, A
C 30	26.2	9.6	201	18	US-10-719-993-34280	Sequence 34280, A
C 31	26.2	9.6	201	18	US-10-719-993-34282	Sequence 34282, A
C 32	26.2	9.6	201	18	US-10-719-993-53459	Sequence 53459, A
C 33	26.2	9.6	201	18	US-10-719-993-53460	Sequence 53460, A
C 34	26.2	9.6	216	18	US-10-653-047-2852	Sequence 2852, Ap
C 35	26	9.5	198	9	US-09-864-761-26345	Sequence 26345, A
C 36	26	9.5	198	15	US-10-054-435-30	Sequence 30, Appl
C 37	26	9.5	201	18	US-10-719-993-16143	Sequence 16143, A
C 38	26	9.5	249	15	US-10-127-032-36	Sequence 36, Appl
C 39	26	9.5	266	18	US-10-425-115-101750	Sequence 101750,
C 40	25.8	9.5	169	9	US-09-867-701-3189	Sequence 3189, Ap
C 41	25.8	9.5	201	18	US-10-719-993-34124	Sequence 34124, A
C 42	25.8	9.5	243	18	US-10-425-115-24398	Sequence 24398, A
C 43	25.8	9.5	261	18	US-10-425-115-79105	Sequence 79105, A
C 44	25.6	9.4	87	14	US-10-083-357-194	Sequence 194, App
C 45	25.6	9.4	201	17	US-10-741-601-19290	Sequence 19290, A

ALIGNMENTS

RESULT 1
US-10-264-049-1744/c
; Sequence 1744, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/185569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1744
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-1744

Query Match	10.9%;	Score 29.8;	DB 16;	Length 192;
Best Local Similarity	58.4%;	Pred. No. 11;		
Matches	52;	Conservative	0;	Mismatches 37; Indels 0; Gaps 0;
QY	63	ACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCGAACCTGCTGGG	122	
Db	128	ACTTTCTTGTAGGCTTCAGACGCACGACCTTGAGGGCAGCAGGACCATCCGCTTTT	69	
QY	123	GTTGGCTCTGGGTCTGGTTGGAGGGCGCA	151	
Db	68	TCTTGTCTAGGGCGGTGGGATGCCGTCA	40	

QY 198 GGCTGTCAGCTGCGCTGAGATTGTCAAGCGGGGTCCAGGCTGCACCAAGCTCACAA 257
| | | | |
Db 110 GGCTGAGGTCACTGCTGCGCCACCATGCGGTGCTCCAGGACTGCACCCCATCTTCAA 169
| | | | |
QY 258 GCTACGTTTCTTC 271
| | | | |
Db 170 GGTACTTATTATC 183
| | | | |

RESULT 13
US-10-741-601-23785/c
; Sequence 23785, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23785
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-23785

Query Match 10.1%; Score 27.6; DB 17; Length 201;
Best Local Similarity 60.8%; Pred. No. 57;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 198 GGCTGTCAGCTGCGCTGAGATTGTCAAGCGGGGTCCAGGCTGCACCAAGCTCACAA 257
| | | | |
Db 92 GGCTGAGGTCACTGCTGCGCCACCATGCGGTGCTCCAGGACTGCACCCCATCTTCAA 33
| | | | |
QY 258 GCTACGTTTCTTC 271
| | | | |
Db 32 GGTACTTATTATC 19
| | | | |

RESULT 14
US-10-641-428-8/c
; Sequence 8, Application US/10641428
; Publication No. US20040121368A1
; GENERAL INFORMATION:
; APPLICANT: Gifu University
; TITLE OF INVENTION: Polynucleotides, polypeptides and method for
; TITLE OF INVENTION: screening for useful dog candidates.
; FILE REFERENCE: P1P2002381US
; CURRENT APPLICATION NUMBER: US/10/641,428
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: JP2002-373006
; PRIOR FILING DATE: 2002-12-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 133
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1)..(133)
; OTHER INFORMATION: Dopamine Receptor D4 Intron 2 (Long form)
US-10-641-428-8

Query Match 10.0%; Score 27.2; DB 17; Length 133;
Best Local Similarity 53.8%; Pred. No. 74;
Matches 56; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 137 GGTGGAGGGCGGCAAGTCTCGGCATGTAGTTCTCAGGTTCTGGCAGGCTGCAGGAA 196
| | | | |
Db 116 GGGCGGCGGGCCGGGGCGGGGCTGGCGAGGGGCGGGGCGGCGGCTGGCGAG 57
| | | | |

QY 197 AGGCTGTACGCTGCGCTGAGATTGTCAAGCGGGGTCCAGGCTGCACCAAGCTCACAGGC 240
| | | | |
Db 56 GGGCGGGTCCGAGAGGGCGGGCCTGGCGGGGGCGGGGGCGGGCGGCGGC 13
| | | | |

RESULT 15
US-10-872-874-287/c
; Sequence 287, Application US/10872874
; Publication No. US20050003419A1
; GENERAL INFORMATION:
; APPLICANT: Breves, Roland
; APPLICANT: Maurer, Karl-Heinz
; APPLICANT: Eck, Jurgen
; APPLICANT: Lorenz, Patrick
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: GLYCOSYL HYDROLASES
; FILE REFERENCE: HENK-0088 / H 5206
; CURRENT APPLICATION NUMBER: US/10/872,874
; CURRENT FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: PCT/EP02/14210
; PRIOR FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: DE 101 63 748.9
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 287
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Synechocystis sp
US-10-872-874-287

Query Match 10.0%; Score 27.2; DB 18; Length 235;
Best Local Similarity 50.0%; Pred. No. 79;
Matches 68; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 115 CTGCTGGGGTGGCTCTGGGTGCGGTGAGGGCGGCAGTGCTCGGCATGTAGTTCTCA 174
| | | | |
Db 205 CTCCTTGGCCCAAGGTGACGACCGAGTCCGACGAGCTTGCCCATCATGCGGTTCTCGGT 146
| | | | |

QY 175 GGTTCCTGGCAGGGCTGCAGGAAAGGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGGGGTC 234
| | | | |
Db 145 GGCGGTGTTCGAGCAGCAGGTGCTGTTGGCCACCGAGCCGTCGGCGAGCCGCTGGTA 86
| | | | |
QY 235 CCAGGCCCTGCACCAGC 250
| | | | |
Db 85 GTAGCCGGGCACGATC 70
| | | | |

Search completed: January 31, 2005, 21:07:51
Job time : 339 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:01:47 ; Search time 2458 Seconds
(without alignments)
4503.816 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaaagtacgttcccttcag 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 55031126

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:*
118: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
119: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:*
120: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:*
121: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:*
122: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:*
123: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:*
124: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:*
125: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	226	82.8	264	13	US-08-856-624-2456	Sequence 2456, Ap
2	226	82.8	264	24	US-09-540-766-8863	Sequence 8863, Ap
3	218	79.9	271	14	US-08-941-869A-2841	Sequence 2841, Ap
4	218	79.9	271	24	US-09-540-208-35269	Sequence 35269, A
5	212	77.7	245	9	US-08-438-571A-2389	Sequence 2389, Ap
6	212	77.7	245	24	US-09-540-233D-126181	Sequence 126181, Ap
7	206	75.5	260	14	US-08-959-395-768	Sequence 768, Ap
8	206	75.5	260	23	US-09-539-334-6888	Sequence 6888, Ap
9	196	71.8	262	13	US-08-812-505-1415	Sequence 1415, Ap
10	196	71.8	262	14	US-08-978-620-1415	Sequence 1415, Ap
11	196	71.8	262	24	US-09-540-229-45100	Sequence 45100, A
12	196	71.8	262	66	US-60-013-529-1415	Sequence 1415, Ap
13	193	70.7	250	14	US-08-951-197-5067	Sequence 5067, Ap
14	193	70.7	250	23	US-09-539-806-30498	Sequence 30498, A
15	193	70.7	250	23	US-09-539-806B-30498	Sequence 30498, A
16	193	70.7	250	67	US-60-027-249-1037	Sequence 1037, Ap
17	192	70.3	250	23	US-09-539-800-16108	Sequence 16108, A
18	192	70.3	250	23	US-09-539-800B-16108	Sequence 16108, A
19	192	70.3	250	23	US-09-539-800C-16108	Sequence 16108, A
20	190	69.6	244	15	US-09-014-441-1305	Sequence 1305, Ap
21	190	69.6	244	24	US-09-540-212A-30093	Sequence 30093, A
22	190	69.6	244	68	US-60-034-841-1305	Sequence 1305, Ap
23	189.6	69.5	264	24	US-09-540-213-18789	Sequence 18789, A
24	185	67.8	239	14	US-08-901-902-73	Sequence 73, Appl
25	185	67.8	239	23	US-09-539-806-30137	Sequence 30137, A
26	185	67.8	239	23	US-09-539-806B-30137	Sequence 30137, A
27	185	67.8	239	67	US-60-023-379-73	Sequence 73, Appl
28	184	67.4	233	13	US-08-879-863-3710	Sequence 3710, Ap
29	184	67.4	233	13	US-08-879-863A-3710	Sequence 3710, Ap
30	184	67.4	233	24	US-09-540-213-12679	Sequence 12679, A
31	183.4	67.2	259	18	US-09-293-657-1609	Sequence 1609, Ap
32	183.4	67.2	259	24	US-09-540-233D-76641	Sequence 76641, A
33	183.4	67.2	259	73	US-60-082-035-1609	Sequence 1609, Ap
34	183	67.0	238	23	US-09-539-334-28833	Sequence 28833, A
35	181	66.3	238	24	US-09-540-229-113556	Sequence 113556, Ap
36	169	61.9	227	9	US-08-413-150-22	Sequence 22, Appl
37	169	61.9	227	24	US-09-540-499-10618	Sequence 10618, A
38	169	61.9	227	24	US-09-540-499B-10618	Sequence 10618, A
39	166	60.8	224	23	US-09-539-800-15324	Sequence 15324, A
40	166	60.8	224	23	US-09-539-800B-15324	Sequence 15324, A
41	166	60.8	224	23	US-09-539-800C-15324	Sequence 15324, A
42	166	60.8	234	14	US-08-923-902-2187	Sequence 2187, Ap
43	166	60.8	234	24	US-09-540-208-7903	Sequence 7903, Ap
44	157	57.5	210	13	US-08-801-504-2052	Sequence 2052, Ap
45	157	57.5	210	24	US-09-540-499-16308	Sequence 16308, A

ALIGNMENTS

RESULT 1
US-08-856-624-2456
; Sequence 2456, Application US/08856624
; GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
NUMBER OF SEQUENCES: 3070
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,624
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017,190
FILING DATE: MAY 13, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,100
FILING DATE: SEPTEMBER 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0174 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 2456:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 1929519
US-08-856-624-2456

Query Match 82.8%; Score 226; DB 13; Length 264;
Best Local Similarity 100.0%; Pred. No. 6.4e-51;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 CCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAAT 107
Db 1 CCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAAT 60
Qy 108 TCGCAACCTGCTGGGGTGGCTCTGGGTGGTGGAGGGCGGAGTGTCTCGGCATGTAGT 167
Db 61 TCGCAACCTGCTGGGGTGGCTCTGGGTGGTGGAGGGCGGAGTGTCTCGGCATGTAGT 120
Qy 168 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTCAGCTCGCTGAGATTGTCAAGCG 227
Db 121 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAAGGCTCAGCTCGCTGAGATTGTCAAGCG 180
Qy 228 GCGGGTCCCAGGCTGCACCCAGCTCACCAAGCTACGTTTCTTTCAG 273
Db 181 GCGGGTCCCAGGCTGCACCCAGCTCACCAAGCTACGTTTCTTTCAG 226

```
US-09-540-766-8863
;
; Sequence 8863, Application US/09540766
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
; FILE REFERENCE: PD-1024 CIP
; CURRENT APPLICATION NUMBER: US/09/540,766
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
; NUMBER OF SEQ ID NOS: 77960
; SOFTWARE: PERL Program
; SEQ ID NO 8863
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00958497
US-09-540-766-8863

Query Match      82.8%; Score 226; DB 24; Length 264;
Best Local Similarity 100.0%; Pred. No. 6.4e-51;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CCCAATGCCCCAGCTACCTCTGATACCCCTGAGATCGGGTCCGAGATGGCAGCAAAAT 107
|
|
|
Db 1 CCCAATGCCCCAGCTACCTCTGATACCCCTGAGATCGGGTCCGAGATGGCAGCAAAAT 60
|
|
|
QY 108 TCGCAACCTGCTGGGGTGGCTCTGGGTCGGTGGAGGGCGGAGTGTGCGCATGTAGT 167
|
|
|
Db 61 TCGCAACCTGCTGGGGTGGCTCTGGGTCGGTGGAGGGCGGAGTGTGCGCATGTAGT 120
|
|
|
QY 168 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGAGATTGTCAAGCG 227
|
|
|
Db 121 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGAGATTGTCAAGCG 180
|
|
|
QY 228 GCGGTCCCGAGGCTGCACAGCTCACCAGCTACGTTTCCTTCAG 273
|
|
|
Db 181 GCGGTCCCGAGGCTGCACAGCTCACCAGCTACGTTTCCTTCAG 226
|
|
|

RESULT 3
US-08-941-869A-2841
; Sequence 2841, Application US/08941869A
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 5486
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
```

```
US-09-540-766-8863
;
; Sequence 8863, Application US/08/941,869A
; FILING DATE: HERewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,839
; FILING DATE: SEPTEMBER 30, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/012,689
; FILING DATE: FEBRUARY 27, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0119-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 2841:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1794787
US-08-941-869A-2841

Query Match      79.9%; Score 218; DB 14; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGGCTTCCCCAATGCCCCAG 60
|
|
|
Db 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGGCTTCCCCAATGCCCCAG 113
|
|
|
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
|
|
|
Db 114 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
|
|
|
QY 121 GGGTTGGCTCTGGGTCGGTGGAGGGCGGAGTGTGCGCATGTAGTGTCTCAGGTTCT 180
|
|
|
Db 174 GGGTTGGCTCTGGGTCGGTGGAGGGCGGAGTGTGCGCATGTAGTGTCTCAGGTTCT 233
|
|
|
QY 181 GGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGCGCTGAGAT 218
|
|
|
Db 234 GGCAGGGCTGCAGGAAGGCTGTGAGTGGCTGCGCTGAGAT 271
|
|
|

RESULT 4
US-09-540-208-35269
; Sequence 35269, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 35269
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```



```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00317250
US-09-540-208-35269

Query Match      79.9%; Score 218; DB 24; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 113
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 114 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 174 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCTCGGCTGAGAT 218
Db 234 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCTCGGCTGAGAT 271
```

```
RESULT 5
US-08-438-571A-2389
; Sequence 2389, Application US/08438571A
; GENERAL INFORMATION:
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Levine, Wendy B.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Freitas, Nicole
; APPLICANT: Jernigan-Kelleher, Colleen
; APPLICANT: Stuart, Susan G.
; APPLICANT: Scott, Randall W.
; APPLICANT: Bills, Pamela Kay
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Akerblom, Ingrid Erika
; TITLE OF INVENTION: POLYNUCLEOTIDES DERIVED FROM THP-1 CELLS
; NUMBER OF SEQUENCES: 5094
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,571A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph.D.
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-001-5 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2389:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: P-033488
US-08-438-571A-2389

Query Match      77.7%; Score 212; DB 9; Length 245;
Best Local Similarity 98.1%; Pred. No. 4.1e-47;
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 25 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 84
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 85 CTACCTCCTGATACCCCTTNANATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 144
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 145 GGGTTGGCTCTGGTTCGGTGGAGGGCGGAGTGGCTCGGCATGTAGTGTCTCAGGTTCT 204
QY 181 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCTCGGCTGAG 216
Db 205 GGCAGGGCTGCAGGAAAGCTGTGAGTGGCTCGGCTGAG 240

RESULT 6
US-09-540-233D-126181
; Sequence 126181, Application US/09540233D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF HEMIC AND IMMUNE SYSTEM TISSUE
; FILE REFERENCE: PD-1030 CIP
; CURRENT APPLICATION NUMBER: US/09/540,233D
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 8/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/435,761
; PRIOR FILING DATE: 1995-05-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 129340
; SOFTWARE: PERL Program
; SEQ ID NO 126181
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00448643
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 103, 105, 178, 231
```

```

; OTHER INFORMATION: a, t, c, g, or other
US-09-540-233D-126181

Query Match      77.7%; Score 212; DB 24; Length 245;
Best Local Similarity 98.1%; Pred. No. 4.1e-47;
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 25 ATGAGGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 84
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 85 CTACCTCCTGATACCCCTTNANATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 144
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGCGTGTAGTGGCATGTAGTGTCTCAGGTTCT 180
Db 145 GGGTTGGCTCTGGGTGGAGGGCGGCGTGTAGTGGCATGTAGTGTCTCAGGTTCT 204
QY 181 GGCAGGGCTGCAGGAAAGGCTGTAGTGGCATGTAGTGGCATGTAGTGTCTCAGGTTCT 216
Db 205 GGCAGGGCTGCAGGAAAGGCTGTAGTGGCATGTAGTGGCATGTAGTGTCTCAGGTTCT 240

RESULT 7
US-08-959-395-768
; Sequence 768, Application US/08959395
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN PANCREATIC ISLET CELLS
; NUMBER OF SEQUENCES: 5789
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,395
; FILING DATE: HEREWITH
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,755
; FILING DATE: OCTOBER 28, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,551
; FILING DATE: DECEMBER 20, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0259 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 768:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 260 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 2071192H2
US-08-959-395-768

Query Match      75.5%; Score 206; DB 14; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 49 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 108
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 109 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 168
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 169 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 228
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCTAG 206
Db 229 GGCAGGGCTGCAGGAAAGGCTGTCTAG 254

RESULT 8
US-09-539-334-6888
; Sequence 6888, Application US/09539334
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF ENDOCRINE SYSTEM TISSUE
; FILE REFERENCE: PD-1026 CIP
; CURRENT APPLICATION NUMBER: US/09/539,334
; CURRENT FILING DATE: 2000-03-30
; "Prior application data removed - refer to PALM or file wrapper"
; NUMBER OF SEQ ID NOS: 38381
; SOFTWARE: PERL Program
; SEQ ID NO 6888
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00887627
US-09-539-334-6888

Query Match      75.5%; Score 206; DB 23; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 60
Db 49 ATGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAG 108
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 120
Db 109 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG 168
QY 121 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 169 GGGTTGGCTCTGGGTGGAGGGCGGCGAGTGTCTCGGCATGTAGTGTCTCAGGTTCT 228
QY 181 GGCAGGGCTGCAGGAAAGGCTGTCTAG 206
Db 229 GGCAGGGCTGCAGGAAAGGCTGTCTAG 254
```

RESULT 9
US-08-812-505-1415
; Sequence 1415, Application US/08812505
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: ASTROCYTOMA
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,505
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/013,529
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-812-505-1415
Query Match 71.8%; Score 196; DB 13; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCCAATGCCCCAG 60
Db 60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 119
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 120 CTACCTCCTGATACCCCTTNANATGCGGGTCCGAGATGGCAGCAAAATNGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGGTTCGGTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 180 GGGTTGGCTCTGGGTTCGGTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 239
QY 181 GGCAGGGCTGCAGGAAAGGCT 201
Db 240 GGCAGGGCTGCAGGAAAGGCT 260

RESULT 10
US-08-978-620-1415
; Sequence 1415, Application US/08978620
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN BRAIN
; NUMBER OF SEQUENCES: 7797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,620
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,151
; FILING DATE: DECEMBER 6, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/812,505
; FILING DATE: MARCH 7, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-978-620-1415
Query Match 71.8%; Score 196; DB 14; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCCAATGCCCCAG 60
Db 60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 119
QY 61 CTACCTCCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 120 CTACCTCCTGATACCCCTTNANATGCGGGTCCGAGATGGCAGCAAAATNGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGGTTCGGTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 180 GGGTTGGCTCTGGGTTCGGTGGAGGCGGCGAGTCTCGGCATGTAGTGTCTCAGGTTCT 239

Qy	181	GGCAGGGCTGCAGGAAGGCT	201
	240	GGCAGGGCTGCAGGAAGGCT	260
Db			

RESULT 11

```

US-09-540-229-45100
; Sequence 45100; Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 45100
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00691274
; NAME/KEY: unsure
; LOCATION: 38, 76, 87, 138, 140, 168
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-45100

```

	Query Match	71.8%	Score 196;	DB 24;	Length 262;
	Best Local Similarity	97.5%;	Pred. No. 9.5e-43;		
	Matches 196;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1	ATGGAGCAC	TACCGGAAAGTGGCTCTGTAGAGCTCCAGGCCTTCCC	AATGCCCCAG	60
Dd	60	ATGGAGCAC	TACCGGANAGTGGCTCTNTAGAGCTCCAGGCCTTCCC	AATGCCCCAG	119
QY	61	CTACCTCCTGATA	ACCCTTGAGATCGGGTCCGAGATGGCAGCAAAAATT	CGCAACCTGCTG	120
Dd	120	CTACCTCCTGATA	ACCCTTNANATCGGGTCCGAGATGGCAGCAAAAATT	NGCAACCTGCTG	179
QY	121	GGGTTGGCTCT	GGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGT	TCTCAGGTTCT	180
Dd	180	GGGTTGGCTCT	GGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGT	TCTCAGGTTCT	239

RESULT 12

US-60-013-529-1415
; Sequence 1415, Application US/60013529
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: ASTROCYTOMA
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/013,529

Query Match	71.8%;	Score 196;	DB 66;	Length 262;
Best Local Similarity	97.5%;	Pred. No. 9.5e-43;		
Matches 196;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
1	ATGGAGCACTACCGAAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG	60		
60	ATGGAGCACTACCGANAGCTGGCTCTNTAGAGCTCCAGCGCCTTCCCAATGCCCCAG	119		
61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGCCAGCAAAATTGCAACCTGCTG	120		
120	CTACCTCCTGATACCCCTTNANATCGGGTCCGAGATGCCAGCAAAATTNGCAACCTGCTG	179		
121	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT	180		
180	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTCGGCATGTAGTGTCTCAGGTTCT	239		
181	GGCAGGGCTGCAGGAAAGGCT	201		
240	GGCAGGGCTGCAGGAAAGGCT	260		

RESULT 13

US-08-951-197-5067
; Sequence 5067, Application US/08951197
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuv, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM HUMAN
; NUMBER OF INVENTION: BREAST
; NUMBER OF SEQUENCES: 5970
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,197
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,249
; FILING DATE: OCTOBER 1, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/826,438
; FILING DATE: MARCH 20, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,285
; FILING DATE: MARCH 20, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0143-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5067:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1997229
; US-08-951-197-5067

Query Match 70.7%; Score 193; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGTCCCAATGCCCGAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGTCCCAATGCCCGAG 117

QY 61 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177

QY 121 GGGTTGGCTCTGGTGGTGGAGGGCGGCACTGCTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178 GGGTTGGCTCTGGTGGTGGAGGGCGGCACTGCTGGCATGTAGTGTCTCAGGTTCT 237

QY 181 GGCAGGGCTGCAG 193
Db ||||||||||||
238 GGCAGGGCTGCAG 250

RESULT 14
US-09-539-806-30498
; Sequence 30498, Application US/09539806
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE

FILE REFERENCE: PD-1027 CIP
CURRENT APPLICATION NUMBER: US/09/539,806
CURRENT FILING DATE: 2000-03-30
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 48372
SOFTWARE: PERL Program
SEQ ID NO 30498
LENGTH: 250
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: hu00555791
US-09-539-806-30498

Query Match 70.7%; Score 193; DB 23; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGTCCCAATGCCCGAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGTCCCAATGCCCGAG 117

QY 61 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 CTACCTCTGATACCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177

QY 121 GGGTTGGCTCTGGTGGTGGAGGGCGGCACTGCTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178 GGGTTGGCTCTGGTGGTGGAGGGCGGCACTGCTGGCATGTAGTGTCTCAGGTTCT 237

QY 181 GGCAGGGCTGCAG 193
Db ||||||||||||
238 GGCAGGGCTGCAG 250

RESULT 15
US-09-539-806B-30498
; Sequence 30498, Application US/09539806B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539,806B
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/706,766
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/004,676
; PRIOR FILING DATE: October 2, 1995
; PRIOR APPLICATION NUMBER: 08/749,515
; PRIOR FILING DATE: November 15, 1996
; PRIOR APPLICATION NUMBER: 60/006,810
; PRIOR FILING DATE: November 15, 1995
; PRIOR APPLICATION NUMBER: 08/822,285
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/013,696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/951,197
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027,249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/826,438
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/016,145
; PRIOR FILING DATE: April 18, 1996
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 48372
; SOFTWARE: PERL Program

SEQ ID NO 30498
LENGTH: 250
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: hu00555791
US-09-539-806B-30498

Query Match 70.7%; Score 193; DB 23; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG	60
Db	58	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCGCCTTCCCCAATGCCCCAG	117
QY	61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG	120
Db	118	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG	177
QY	121	GGGTGGCTCTGGTGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT	180
Db	178	GGGTGGCTCTGGTGGTGGAGGGCGGCGAGTGTGGCATGTAGTGTCTCAGGTTCT	237
QY	181	GGCAGGGCTGCAG	193
Db	238	GGCAGGGCTGCAG	250

Search completed: January 31, 2005, 20:58:10
Job time : 2461 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:09:11 ; Search time 74 Seconds
(without alignments)
2622.235 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaagctacgtttccttcag 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1158410

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	9.9	264	4	US-09-370-838-7
2	27	9.9	264	4	US-09-854-133-7
C 3	26	9.5	92	3	US-09-511-625B-45
4	26	9.5	198	3	US-09-079-812E-30
C 5	25.8	9.5	165	3	US-08-905-223-154
C 6	25.6	9.4	212	4	US-09-313-294A-2448
C 7	25.6	9.4	261	4	US-09-489-039A-1139
8	25.4	9.3	194	4	US-09-513-999C-22180
9	25.4	9.3	235	4	US-09-016-434-661
C 10	25	9.2	211	4	US-09-513-999C-14087
C 11	25	9.2	255	4	US-09-513-999C-36531
C 12	24.6	9.0	215	4	US-09-513-999C-2518
13	24.6	9.0	272	4	US-09-016-434-855
C 14	24.4	8.9	95	4	US-09-513-999C-15097
C 15	24.4	8.9	261	4	US-09-489-039A-2953
C 16	24.2	8.9	163	4	US-09-621-976-7986
C 17	24.2	8.9	180	4	US-09-513-999C-22891
18	24	8.8	140	4	US-09-513-999C-20498
19	24	8.8	181	4	US-09-513-999C-20847
20	23.8	8.7	137	4	US-09-513-999C-9178
21	23.8	8.7	231	4	US-09-252-991A-2750
C 22	23.6	8.6	139	4	US-09-513-999C-17624
C 23	23.6	8.6	183	4	US-09-513-999C-15131
24	23.6	8.6	186	3	US-09-188-930-95
25	23.6	8.6	186	4	US-09-312-283C-95
C 26	23.6	8.6	234	1	US-08-420-235B-6
C 27	23.6	8.6	234	3	US-08-793-624-6

C	28	23.6	8.6	234	5	PCT-US95-10194-6	Sequence 6, Appli
	29	23.4	8.6	167	4	US-09-621-976-19219	Sequence 19219, A
	30	23.2	8.5	160	4	US-09-513-999C-12509	Sequence 12509, A
C	31	23.2	8.5	169	1	US-08-519-777-29	Sequence 29, Appl
C	32	23.2	8.5	169	1	US-08-742-035-29	Sequence 29, Appl
C	33	23.2	8.5	169	1	US-08-777-019-29	Sequence 29, Appl
C	34	23.2	8.5	169	2	US-08-777-143-29	Sequence 29, Appl
C	35	23.2	8.5	169	3	US-08-775-414-29	Sequence 29, Appl
C	36	23.2	8.5	169	3	US-08-931-858E-29	Sequence 29, Appl
C	37	23.2	8.5	169	3	US-08-981-739-29	Sequence 29, Appl
C	38	23.2	8.5	169	4	US-09-128-026-29	Sequence 29, Appl
C	39	23.2	8.5	169	4	US-09-220-616-29	Sequence 29, Appl
C	40	23.2	8.5	169	4	US-09-220-527-29	Sequence 29, Appl
C	41	23.2	8.5	169	4	US-09-220-407-29	Sequence 29, Appl
C	42	23.2	8.5	169	4	US-09-219-952-29	Sequence 29, Appl
	43	23.2	8.5	192	4	US-09-489-039A-7126	Sequence 7126, Ap
C	44	23.2	8.5	246	4	US-09-489-039A-7074	Sequence 7074, Ap
	45	23.2	8.5	251	4	US-09-313-294A-3512	Sequence 3512, Ap

ALIGNMENTS

RESULT 1

US-09-370-838-7
; Sequence 7, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(264)
; OTHER INFORMATION: n = A,T,C or G
US-09-370-838-7

Query Match 9.9%; Score 27; DB 4; Length 264;
Best Local Similarity 54.3%; Pred. No. 42;
Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

Qy	175	GGTTCTGGCAGGGCTGCAGGAAGGCTGTCTCAGCTGCGCTGAGATTGTCAAGCGGGGTC	234
Db	12	GGATGGGGCACGGGAGCACAGATMGACTTTAACTGCCCCCAGCTTTCMAGGAAGGATT	71
Qy	235	CCAGGCTGCACAGCTCACCAGCTCACCAGCTACGTTTCC	268
Db	72	ACAGGCGTGAGCCACTGCGCCCGGCTCTTCTCC	105

RESULT 2

US-09-854-133-7
; Sequence 7, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.


```

; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(264)
; OTHER INFORMATION: n = A,T,C or G
US-09-854-133-7

Query Match          9.9%; Score 27; DB 4; Length 264;
Best Local Similarity 54.3%; Pred. No. 42;
Matches 51; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 175 GGTCTGCGAGGCTGCAGGAAAGCTGTGCTGCGCTGCGTGAAGGCGGGTCC 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12  GGATGGGCGACGGGAGCACAGATMGACTTTAACTGCCCCCAGCTTTCMAGGAAAGGATT 71

QY 235 CCAGGCCTGCACCAAGCTCACCAGCTACGTTTCC 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72  ACAGGCGTGAGCCACTGGCCCGGCTCTTCTCC 105
```

```

RESULT 3
US-09-511-625B-45/c
; Sequence 45, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
; APPLICANT: LaRoche, William J.
; APPLICANT: Patel, Bharvin
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: ATTENUATED AND DOMINANT NEGATIVE VARIANT
; FILE REFERENCE: 14014.0300u1
; CURRENT APPLICATION NUMBER: US/09/511,625B
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: PCT/US98/17821
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 60/070,397
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 60/056,075
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/No. 6368828e -
; OTHER INFORMATION: synthetic construct
US-09-511-625B-45

Query Match          9.5%; Score 26; DB 3; Length 92;
Best Local Similarity 57.3%; Pred. No. 59;
Matches 47; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 179 CTGGCAGGCTGCAGGAAAGCTGTGCTGCGCTGAGATTGTCAAGCGGGTCCCAG 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92  CTGGCCGGGATGAAGGAGGATGGGGCATGGGCAGTGATGAGGTGACTGGTGTAT 33

QY 239 GCCTGCACCAAGCTCACCAGCT 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32  GGCTGCTCAAACTACCCAGGTT 11
```

```

RESULT 4
US-09-079-812E-30
; Sequence 30, Application US/09079812E
; Patent No. 6340575
; GENERAL INFORMATION:
; APPLICANT: Bollag, Gideon
; APPLICANT: Crompton, Anne
; APPLICANT: No. 6340575th, Anne
; APPLICANT: Sharma, Sanju
; APPLICANT: Roscoe, William
; TITLE OF INVENTION: Methods and Compositions for Treating Abnormal Cell
; TITLE OF INVENTION: Growth Related to Unwanted Guanine Nucleotide Exchange
; TITLE OF INVENTION: Factor Activity
; FILE REFERENCE: 1028-US
; CURRENT APPLICATION NUMBER: US/09/079,812E
; CURRENT FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/049,879
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Glu-Epitope Tag
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement (1)..(198)
US-09-079-812E-30
```

```

Query Match          9.5%; Score 26; DB 3; Length 198;
Best Local Similarity 57.3%; Pred. No. 77;
Matches 47; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 182 GCAGGGCTGCAGGAAAGCTGTGCTGCGCTGCGTGAAGCGGGTCCCAGGCC 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85  GCAGACCCTATGAGCTGTGCGCGCTGCTTCAGCGCGCTGCGCGCTGGACCGAC 144

QY 242 TGCACCAAGCTCACCAGCTACG 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 AGTCCACCTCGGAAAGCGACG 166
```

```

RESULT 5
US-08-905-223-154/c
; Sequence 154, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 154:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 46..153
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 3.8
OTHER INFORMATION: seq GTDSLFLPPCPC/CP
US-08-905-223-154

Query Match 9.5%; Score 25.8; DB 3; Length 165;
Best Local Similarity 63.9%; Pred. No. 84;
Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 173 CAGGTTCTGGCAGGCTGCAGGAAGGCTGTCTCAGCTGCGCTGAGATTGCAAGCGCGGG 232
|||||
Db 152 CAGGGCAGGTTGGTAGGAAGGAAGGAGTCAGTTCCTCACTGAACTCCCAACTGAGGA 93
|||||

QY 233 T 233
|
Db 92 T 92

RESULT 6

US-09-313-294A-2448/c
Sequence 2448, Application US/09313294A
Patent No. 6476212

GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2448
LENGTH: 212
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700552684H1
US-09-313-294A-2448

Query Match 9.4%; Score 25.6; DB 4; Length 212;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 129 TCTGGTCTGGAGGGCGGAGTCTCGGCATGTAGTGTCTCAGGTTCTGGCAGGGC 188
|||||
Db 130 TCTGCCAGGTGAGCGGGCACCGGGCTGGGCATGTACAGAGGTTAGGGTTGGCTTGGA 71
|||||

QY 189 TGCA 192
||
Db 70 GCCA 67

RESULT 7

US-09-489-039A-1139/c
Sequence 1139, Application US/09489039A

Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 1139
LENGTH: 261
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1139

Query Match 9.4%; Score 25.6; DB 4; Length 261;
Best Local Similarity 57.5%; Pred. No. 1.1e+02;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 127 GCTCTGGTCTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCTGGCAGG 186
|||||
Db 209 GCTGGGCTTGGCTGGCTGCCGGCGTGATAGCCCTCGCGTTGTATGCATAGTCCCGGT 150
|||||

QY 187 GCTGCAGGAAAGGCTGTCTAG 206
|||||
Db 149 GCTGCAGGGGACGCTGGCCG 130
|||||

RESULT 8

US-09-513-999C-22180
Sequence 22180, Application US/09513999C
Patent No. 6783961

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961

FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 22180
LENGTH: 194
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 139
OTHER INFORMATION: r=a or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: 194
OTHER INFORMATION: r=a or g
US-09-513-999C-22180

Query Match 9.3%; Score 25.4; DB 4; Length 194;
Best Local Similarity 52.3%; Pred. No. 1.2e+02;
Matches 56; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 128 CTCTGGGTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTCTCAGGTTCTGGCAGGG 187
|||||
Db 46 CACTGTGCCCCCGGCGAGCACCAGGACTCTGCTGGTTAGGGCTGCGCGGTGAGACAGGG 105
|||||

QY 188 CTGCAGGAAAGGCTGTCTCAGCTCGGCTGAGATTGTCAAGCGGCGGCTC 234
|||||
Db 106 CGGCCACCTGGTACGTGCGCTGCTGATGCGCTTGRCTGCGGCCAGTC 152
|||||


```
; OTHER INFORMATION: w=a or t
US-09-513-999C-36531

Query Match          9.2%; Score 25; DB 4; Length 255;
Best Local Similarity 54.1%; Pred. No. 1.7e+02;
Matches 46; Conservative 2; Mismatches 37; Indels 0; Gaps 0;

QY 132 GGGTCGGTTGGAGGCGGCGAGTCTCGGCATGTACTGTCTCTCAGGTTCTGGCAGGCTGC 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 229 GGTTCGTGTACCGGCGGCAACGTGWCCTGGAGTTGTACAGATKACAGCCTGTGGGC 170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 192 AGGAAAGGCTGTACGCTGCGCTGAG 216
    ||||| ||||| ||||| ||||| |||||
Db 169 AGCGCTGGDCGGCGGCGCYGTAG 145
    ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-513-999C-2518/c
; Sequence 2518, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2518
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..213
US-09-513-999C-2518

Query Match          9.0%; Score 24.6; DB 4; Length 215;
Best Local Similarity 70.2%; Pred. No. 2.2e+02;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 219 TGTCAAGCGGCGGTCCAGGCGCTGCACCGCTCACCAAGCTACGTT 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 TCTCAAGCAGCTGGGACGAGGCGCGCCACCCACCCAGCCCAATT 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-016-434-855
; Sequence 855, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 855:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PGANNOT01
; CLONE: 619240
; US-09-016-434-855

Query Match          9.0%; Score 24.6; DB 4; Length 272;
Best Local Similarity 48.3%; Pred. No. 2.3e+02;
Matches 69; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 115 CTGCTGGGTTGGCTCTGGGTCGGTTGGAGGGCGGCGAGTCTCGGCATGTAGTTCTCA 174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 CTGGCCTGGCTGGCGCTGGGCGAGCATGTGCCTCTTCATCGCCGGCTTTCGGTGGGCTGG 68
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 175 GGTTCCTGGCAGGGCTGCAGGAAAGGCTGTGTCAGCTGCGCTGAGATTGTCAAGCGGCGGTC 234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 GGGCCCATCCCTGGCTCCTCATGTGTCAGAGATCTTCCCTCTGCATGTCAAGGGCGTGGCG 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 235 CCAGGCTGCACCGCTCACCA 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 ACAGGCATCTGCGTCTCACCA 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-513-999C-15097/c
; Sequence 15097, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15097
; LENGTH: 95
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-513-999C-15097

Query Match          8.9%; Score 24.4; DB 4; Length 95;
Best Local Similarity 56.1%; Pred. No. 1.9e+02;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 158 GGCATGTAGTTCTCAGGTTCTGGCAGGCTGCAGGAAAGGCTGTAGTCGCTGAGA 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 GTCATTTTCACTTCTTTTGTGGTGAATGTCGTCAAGGAGGAGGAGGAGGAGGAGGAGG 29
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


QY 218 TTGTCAAGCGGCGGTCCAGG 239
Db 28 ATGTGTACCTGCAGGTCACAGG 7

RESULT 15
US-09-489-039A-2953/c
; Sequence 2953, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2953
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2953

Query Match 8.9%; Score 24.4; DB 4; Length 261;
Best Local Similarity. 47.4%; Pred. No. 2.7e+02;
Matches 73; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 24 CTCTGTAGAGCTCCAGCGCCTTCCCCAATGCCCCAGCTACCTCCTGATACCCCTTGAGAT 83
Db 187 CACCGGCGATCACCCAGCGAGCTGGCGAAATCCCTGCTCGCGCAGGGTGTGGCGGCAT 128
QY 84 GCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGTGGGGTTGGCTCTGGGTCGGTTGGA 143
Db 127 CCGGCTGTGGTCGACGTTGACCATTTCAAAGGCCAGTCCCCGGCTCTCCATGGCTCGCC 68
QY 144 GGGCGGCAGTGTCTGGCATGTAGTGTTCACAGT 177
Db 67 GGGTGGCATGGCACTGCACGCAGTCATTTTCGAGT 34

Search completed: January 31, 2005, 20:59:31
Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 20:58:18 ; Search time 1845 Seconds
(without alignments)
3152.647 Million cell updates/sec

Title: US-10-057-813-13_COPY_253_375
Perfect score: 123
Sequence: 1 accaagctacgtttccttca.....tgctgctcagccgggacccc 123

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 2299606

Minimum DB seq length: 0
Maximum DB seq length: 123

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25.8	21.0	97	6	CQ461900 Sequence
C 2	24.4	19.8	100	6	AX998871 Sequence
C 3	23.8	19.3	101	11	BV185660 sqm14870
C 4	23.4	19.0	89	4	AJ629187 Sus scrofa
C 5	23.4	19.0	100	9	S70266 TAP1 (TAP1-S70266)
C 6	23.4	19.0	100	9	S70274S1 (TAP1-S70274)
C 7	23.4	19.0	101	9	HSITGAD02 Human beta-40275
C 8	23	18.7	89	6	AX899650 Sequence
C 9	23	18.7	89	6	BD035183 Sequence
C 10	23	18.7	121	6	AX262946 Sequence
C 11	23	18.7	121	6	AX262947 Sequence
C 12	22.6	18.4	90	6	A69898 Sequence 27
C 13	22.6	18.4	90	6	BD022719 Mirror pl
C 14	22.2	18.0	117	6	CQ665784 Sequence
C 15	22	17.9	106	3	PCREP2D Z54253 P.clarkii m
C 16	21.8	17.7	75	3	AF400169 Eisenia a
C 17	21.8	17.7	100	6	AX995957 Sequence
C 18	21.8	17.7	100	9	S70256S1 (TAP1-S70256)
C 19	21.6	17.6	75	10	S60901 TCR V beta

20	21.6	17.6	81	14	AF462999	AF462999 Hepatitis G43380 WIAF-2096-S
21	21.6	17.6	120	11	G43380	AX899351 Sequence
C 22	21.4	17.4	119	6	AX899351	BD034884 Sequence
C 23	21.4	17.4	119	6	BD034884	AX388104 Sequence
24	21.2	17.2	91	6	AX388104	AX998870 Sequence
C 25	21.2	17.2	100	6	AX998870	AX042427 Sequence
C 26	21.2	17.2	114	6	AX042427	AX262818 Sequence
C 27	21.2	17.2	121	6	AX262818	AX262819 Sequence
C 28	21.2	17.2	121	6	AX262819	CQ557469 Sequence
C 29	21	17.1	65	6	CQ557469	AF100963 Gorilla g
30	21	17.1	76	9	GGABOB3	AB076772 Homo sapi
31	21	17.1	81	9	AB076772	A69901 Sequence 30
C 32	21	17.1	89	6	A69901	BD022722 Mirror pl
C 33	21	17.1	89	6	BD022722	AB076015 TT virus
34	21	17.1	89	14	AB076015	AX990398 Sequence
C 35	21	17.1	100	6	AX990398	AY302634 Actus tri
C 36	21	17.1	100	9	AY302634S01	AY189210 Human ent
C 37	21	17.1	115	14	AY189210	AY189213 Human ent
C 38	21	17.1	115	14	AY189213	AY189214 Human ent
C 39	21	17.1	116	14	AY189214	AY189216 Human ent
C 40	21	17.1	116	14	AY189216	U71276 Human mutan
41	21	17.1	120	9	HSU71276	U71277 Human mutan
42	21	17.1	120	9	HSU71277	CQ562156 Sequence
43	20.8	16.9	60	6	CQ562156	CQ562325 Sequence
44	20.8	16.9	60	6	CQ562325	S79984 TCR V beta
C 45	20.8	16.9	73	10	S79984	

ALIGNMENTS

RESULT 1					
CQ461900/c					
LOCUS	CQ461900	97 bp	DNA	linear	PAT 30-JAN-2004
DEFINITION	Sequence 4678 from Patent WO0192581.				
ACCESSION	CQ461900				
VERSION	CQ461900.1	GI:41427519			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Algate,P.A., Harlocker,S.L. and Jones,R.				
TITLE	Compositions and methods for the therapy and diagnosis of ovariancancer				
JOURNAL	Patent: WO 0192581-A 4678 06-DEC-2001;				
	CORIXA CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1..97				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Query Match	21.0%;	Score 25.8;	DB 6;	Length 97;
Best local Similarity	56.5%;	Pred. No. 7.2e+03;		
Matches	48;	Conservative	0;	Mismatches 37;
Indels	0;	Gaps	0;	
QY	39	CCCAGCCTACCTGACACAGGGCTAGACCCCTCAGTCAGTGGCGCCATGTGCCTGCAGT	98	
Db	97	CCCCCGCGGCTGCTTCTCTTCTCCGCGCCACTGCCAACCTCCAGTCGCGCCGCGAGA	38	
QY	99	GTGGGTGCTGCTCAGCCGGGACCCC	123	
Db	37	ACCTCCGCTGCTGCTGCGCGGACCCC	13	

RESULT 2

AX998871	AX998871	100 bp	DNA	linear	PAT 16-JAN-2004
LOCUS	Sequence 10334 from Patent EPI260592.				
DEFINITION					
ACCESSION	AX998871				

Qy	2	CCAAGCTACGTTTCTTCAGACTGAGGACAGCTGGTCCCGAGCTCACCTGACACAGGGC	61
Db	29	CCAGCCCTCTCTCCCTCAGACACAGGGGTCCAAAGCCCTCTCTCCCGAGACC	88
Qy	62	TAGACCCCC	70
Db	89	CAGAGGTCC	97
RESULT 4			
LOCUS	AJ629187/c		
DEFINITION	Sus scrofa partial NFYA gene for nuclear transcription factor Y		
ACCESSION	AJ629187		
VERSION	AJ629187.1	GI:44903520	
KEYWORDS	NFYA gene; nuclear transcription factor Y alpha.		
SOURCE	Sus scrofa (pig)		
ORGANISM	Sus scrofa		
REFERENCE	1		
AUTHORS	Barbosa, A., Demeure, O., Urien, C., Milan, D., Chardon, P. and Renard, C.		
TITLE	A physical map of large segments of pig chromosome 7ql.1-q1.4: comparative analysis with human chromosome 6p21		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 89)		
AUTHORS	Barbosa, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-FEB-2004) Barbosa A., Animal Genetic, Institut National de Recherche Agronom., LREG, INRA/CEA, domaine de Vilvert, 78352 Jouy-en-Josas cedex, FRANCE		
FEATURES	Location/Qualifiers		
source	1..89		
	/organism="Sus scrofa"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9823"		
	/chromosome="SSC7ql1"		
	/haplotype="H01"		
	/clone_lib="SBAB genomic BAC library"		
gene	74..89		
	/gene="NFYA"		
CDS	74..>89		
	/gene="NFYA"		
	/codon_start=1		
	/product="nuclear transcription factor Y alpha"		
	/protein_id="CAF32808.1"		
	/db_xref="GI:44903521"		
	/translation="MWSIQ"		
ORIGIN			
	Query Match	19.0%;	Score 23.4; DB 4; Length 89;
	Best Local Similarity	60.0%;	Pred. No. 3.7e+04;
	Matches	39; Conservative	0; Mismatches 26; Indels 0; Gaps 0;
Qy	44	CCTCACCTGACACAGGGCTAGACCCCTCACAGTCGCGCCGATGTGCTGAGTGTGG	103
Db	76	CATTGCCTGCCACTGTGTAGTGTACACCGTGACAGTTCCTTGCCCACTGCTGTGTGG	17
Qy	104	TGCTG	108
Db	16	TGTTG	12
RESULT 5			
LOCUS	S70266/c		
DEFINITION	TAP1 (TAP1-B) {ATP binding region} [human, Caucasoïds, mRNA		
ACCESSION	S70266		
VERSION	S70266.1	GI:546540	

VERSION	AX99871.1	GI:41005217	
KEYWORDS	Escherichia coli		
SOURCE	Escherichia coli		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.		
REFERENCE	1		
AUTHORS	Donner, H., Drescher, B., Huber, A. and Weber, J.		
TITLE	Biochip		
JOURNAL	Patent: EP 1260592-A 10334 27-NOV-2002;		
FEATURES	Location/Qualifiers		
source	1..100		
	/organism="Escherichia coli"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:562"		
	/note="mrec b3251 U00096 complement(3396512__3397615)"		
ORIGIN			
	Query Match	19.8%;	Score 24.4; DB 6; Length 100;
	Best Local Similarity	58.1%;	Pred. No. 1.8e+04;
	Matches	43; Conservative	0; Mismatches 31; Indels 0; Gaps 0;
Qy	5	AGCTACGTTTCTTCAGACTGAGGACAGCTGGTCCCGAGCTCACCTGACACAGGGCTAG	64
Db	27	AGCTACTGGCTGTGTCGGCGACAGAGCTGTCTAATCGCTCTCCACAAAGGGCTAC	86
Qy	65	ACCCCTCACAGTG	78
Db	87	GCCGCGCAAGTG	100
RESULT 3			
LOCUS	BV185660		
DEFINITION	sqnm148706 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.		
ACCESSION	BV185660		
VERSION	BV185660.1	GI:48024577	
KEYWORDS	STS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 101)		
JOURNAL	Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M., Cantor, C.R. and Braun, A.		
COMMENT	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions		
	Genome Res. (2004) In press		
	Contact: Andreas Braun		
	Pharmaceuticals division		
	Sequenom, Inc.		
	3595 John Hopkins Court, San Diego, CA 92121, USA		
	Tel: 18582029018		
	Fax: 18582029020		
	Email: abraun@sequenom.com		
	Primer A: No primer sequence submitted		
	Primer B: No primer sequence submitted		
	STS size: 101.		
FEATURES	Location/Qualifiers		
source	1..101		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/clone_lib="Human DNA (Sequenom)"		
	<1..>101		
STS			
ORIGIN			
	Query Match	19.3%;	Score 23.8; DB 11; Length 101;
	Best Local Similarity	58.0%;	Pred. No. 2.8e+04;
	Matches	40; Conservative	1; Mismatches 28; Indels 0; Gaps 0;

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Szafer,F., Oksenberg,J.R. and Steinman,L.
TITLE New allelic polymorphisms in TAP genes
JOURNAL Immunogenetics 39 (5), 374 (1994)
MEDLINE 94222452
PUBMED 8168860
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 148273] from the original journal article.
FEATURES
source
1..100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..100
/gene="TAP1"
/allele="TAP1-B"
1..100
/gene="TAP1"
/note="Unknown"
/codon_start=1
/protein_id="AAD14055.1"
/db_xref="GI:4261755"
/translation="GYDTEVGEAGSQLSGGQRQAVALARALIRKPCV"
ORIGIN
Query.Match 19.0%; Score 23.4; DB 9; Length 100;
Best Local Similarity 73.2%; Pred. No. 3.6e+04;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 28 GACAGCTGGTCCAGCCTCACCTGACACAGGGCTAGACCC 68
|||||
Db 41 GACAGCTGGTCCAGCCTCGCCTACCTCTGTGCATAGCC 1
RESULT 6
S70274S1/c 100 bp mRNA linear PRI 23-SEP-1994
LOCUS
DEFINITION TAP1 (TAP1-C) {ATP binding and 3' regions} [human, Caucasoids, mRNA
Partial, 100 nt, segment 1 of 2].
ACCESSION S70274
VERSION S70274.1 GI:546541
KEYWORDS
SEGMENT 1 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Szafer,F., Oksenberg,J.R. and Steinman,L.
TITLE New allelic polymorphisms in TAP genes
JOURNAL Immunogenetics 39 (5), 374 (1994)
MEDLINE 94222452
PUBMED 8168860
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 148274] from the original journal article.
FEATURES
source
1..100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..100
/gene="TAP1"
/allele="TAP1-C"
1..100
/gene="TAP1"
/note="Unknown"
gene
CDS

/codon_start=1
/protein_id="AAD14056.1"
/db_xref="GI:4261756"
/translation="GYDTEVGEAGSQLSGGQQQAVALARALIRKPCV"
ORIGIN
Query.Match 19.0%; Score 23.4; DB 9; Length 100;
Best Local Similarity 73.2%; Pred. No. 3.6e+04;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 28 GACAGCTGGTCCAGCCTCACCTGACACAGGGCTAGACCC 68
|||||
Db 41 GACAGCTGGTCCAGCCTCGCCTACCTCTGTGCATAGCC 1
RESULT 7
HSITGAD02 101 bp DNA linear PRI 18-OCT-1996
LOCUS
DEFINITION Human beta-2 integrin alphaD subunit (ITGAD) gene, partial exon 16,
and partial cds.
ACCESSION U40275
VERSION U40275.1 GI:1173860
KEYWORDS
SEGMENT 2 of 6
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Wong,D.A., Davis,E.M., LeBeau,M. and Springer,T.A.
TITLE Cloning and chromosomal localization of a novel gene-encoding a
human beta 2-integrin alpha subunit
JOURNAL Gene 171 (2), 291-294 (1996)
MEDLINE 96257236
PUBMED 8666289
REFERENCE
AUTHORS Wong,D.A.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) Dennis A. Wong, McMaster University,
Medicine, 1200 Main St. West, Hamilton, On L8N 3Z5, Canada
FEATURES
source
1..101
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="alpha W.1"
/clone_lib="pWE15 Genomic"
<1..4
/gene="ITGAD"
<5..>101
/gene="ITGAD"
/codon_start=1
/product="beta-2 integrin alphaD subunit"
/protein_id="AAB60635.1"
/db_xref="GI:1173867"
/translation="RIASSQLSPRLQYFGQALSGGQDLTQDGLMDL"
5..>101
/gene="ITGAD"
/number=16
intron
CDS
exon
ORIGIN
Query.Match 19.0%; Score 23.4; DB 9; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.6e+04;
Matches 39; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 41 CAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGGCCATGTGCCTGCAGTGT 100
|||||
Db 91 CAGTCCATCCTGGGTGAGGTCCTGACCCCCACTCAGCGCCTGTCCAAATACTGCAGCCT 32
GGGTG 105
GGGG 27

AUTHORS	Pekhletski, R., Cooper, R.L., Hampson, D.R. and Atwood, H.L.
TITLE	Expression profiling of mRNA obtained from single identified crustacean motor neurons: determination of specificity of hybridization

JOURNAL
Unpublished

REFERENCE 3 (bases 1 to 106)

AUTHORS Pekhletski, R.

TITLE	Direct Submission
1. <u>THE</u>	
2. <u>THE</u>	
3. <u>THE</u>	
4. <u>THE</u>	
5. <u>THE</u>	
6. <u>THE</u>	
7. <u>THE</u>	
8. <u>THE</u>	
9. <u>THE</u>	
10. <u>THE</u>	
11. <u>THE</u>	
12. <u>THE</u>	
13. <u>THE</u>	
14. <u>THE</u>	
15. <u>THE</u>	
16. <u>THE</u>	
17. <u>THE</u>	
18. <u>THE</u>	
19. <u>THE</u>	
20. <u>THE</u>	
21. <u>THE</u>	
22. <u>THE</u>	
23. <u>THE</u>	
24. <u>THE</u>	
25. <u>THE</u>	
26. <u>THE</u>	
27. <u>THE</u>	
28. <u>THE</u>	
29. <u>THE</u>	
30. <u>THE</u>	
31. <u>THE</u>	
32. <u>THE</u>	
33. <u>THE</u>	
34. <u>THE</u>	
35. <u>THE</u>	
36. <u>THE</u>	
37. <u>THE</u>	
38. <u>THE</u>	
39. <u>THE</u>	
40. <u>THE</u>	
41. <u>THE</u>	
42. <u>THE</u>	
43. <u>THE</u>	
44. <u>THE</u>	
45. <u>THE</u>	
46. <u>THE</u>	
47. <u>THE</u>	
48. <u>THE</u>	
49. <u>THE</u>	
50. <u>THE</u>	
51. <u>THE</u>	
52. <u>THE</u>	
53. <u>THE</u>	
54. <u>THE</u>	
55. <u>THE</u>	
56. <u>THE</u>	
57. <u>THE</u>	
58. <u>THE</u>	
59. <u>THE</u>	
60. <u>THE</u>	
61. <u>THE</u>	
62. <u>THE</u>	
63. <u>THE</u>	
64. <u>THE</u>	
65. <u>THE</u>	
66. <u>THE</u>	
67. <u>THE</u>	
68. <u>THE</u>	
69. <u>THE</u>	
70. <u>THE</u>	
71. <u>THE</u>	
72. <u>THE</u>	
73. <u>THE</u>	
74. <u>THE</u>	
75. <u>THE</u>	
76. <u>THE</u>	
77. <u>THE</u>	
78. <u>THE</u>	
79. <u>THE</u>	
80. <u>THE</u>	
81. <u>THE</u>	
82. <u>THE</u>	
83. <u>THE</u>	
84. <u>THE</u>	
85. <u>THE</u>	
86. <u>THE</u>	
87. <u>THE</u>	
88. <u>THE</u>	
89. <u>THE</u>	
90. <u>THE</u>	
91. <u>THE</u>	
92. <u>THE</u>	
93. <u>THE</u>	
94. <u>THE</u>	
95. <u>THE</u>	
96. <u>THE</u>	
97. <u>THE</u>	
98. <u>THE</u>	
99. <u>THE</u>	
100. <u>THE</u>	

JOURNAL Submitted (28-SEP-1995) Pekhletski R., University of Toronto, Faculty of Pharmacy, 19 Russel Street, Toronto, Ontario, Canada, M5S 2S2

Location/Qualifiers

```
source
1. .106
```

/organism="Procambarus clarkii"

/mol type="mRNA"

```
/db xref="taxon:6728"
```

```

/cell_type="common excitor motor neuron of deep abdominal
extensor"

```

```

/tissue type="nervous system"

```

/dev sta

misc_feature

```

/note="ac-rich sequence
clone only partially sequenced (region proximal to the T7
primer)"

```

ORIGIN

Query Match	17.9%	Score 22;	DB 3;	Length 106;
Best Local Similarity	55.1%	Pred. No. 9.2e+04;		
Matches 43; Conservative	0;	Mismatches 35;	Indels 0;	Gaps 0;

Qy	19	CAGACTGAGGACAGCTGGGTCCCGCCTCACCTGCACACAGGGTAGACCCCCTCACAGTG	78
Db	4	CAGCATCACCCACCCCAACACCCGGCACCAACCACACCCCAACACCCACCCACAGCC	63

QY 79 CGCGCCATGTGCCCTGCA 96

[illegible]

64 CAACCCAGTACCACCA 81

Search completed: January 31, 2005, 21:43:20
Job time : 1848 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 20:59:38 ; Search time 1963 Seconds
(without alignments)
2283.284 Million cell updates/sec

Title: US-10-057-813-13_COPY_253_375
Perfect score: 123
Sequence: 1 accaagctacgttccttca.....tgctgctcagccgggacccc 123

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 1226348

Minimum DB seq length: 0
Maximum DB seq length: 123

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25.8	21.0	97	1 AA430669	AA430669 zw26d03.8
C 2	25.8	21.0	117	2 BF171548	BF171548 PCL2912 M
C 3	25.2	20.5	118	5 BY074656	BY074656 BY074656
4	24.8	20.2	120	1 AV631484	AV631484 AV631484
5	24.6	20.0	114	1 AA092420	AA092420 119365.se
C 6	24.4	19.8	112	8 AZ883677	AZ883677 RPCI-23-1
7	24.2	19.7	115	9 LBAF042F12	BX545265 Leishmani
8	24	19.5	52	1 AI262761	AI262761 QK06e12.x
C 9	24	19.5	122	5 BQ972671	BQ972671 QHD4a07.y
C 10	23.8	19.3	90	1 AA079036	AA079036 zm94h12.r
C 11	23.8	19.3	110	7 N56039	N56039 J6239F Huma
C 12	23.2	18.9	116	8 CC374455	CC374455 PUHDM06TD
C 13	23.2	18.9	120	4 BJ504782	BJ504782 BJ504782
14	23	18.7	113	2 BE165373	BE165373 QV1-HT047
C 15	23	18.7	117	4 BJ472794	BJ472794 BJ472794
16	22.8	18.5	115	9 CC685407	CC685407 OGOAH39TV
17	22.8	18.5	117	2 BF291718	BF291718 WHE2205 B
C 18	22.6	18.4	115	9 CG709891	CG709891 1119015C1
19	22.6	18.4	122	9 CC886444	CC886444 SALK_1486
20	22.4	18.2	101	1 AI208489	AI208489 q955904.x
21	22.4	18.2	104	8 AZ629277	AZ629277 1M0482D12
C 22	22.4	18.2	121	6 CB435325	CB435325 615414 MA
23	22.4	18.2	122	4 BM312764	BM312764 ig78908.y
C 24	22.4	18.2	123	2 BF956593	BF956593 PM1-NN120

25	22.4	18.2	123	2	BE171397	BE171397 RC1-HT054
26	22.2	18.0	79	6	CB298729	CB298729 220024 re
C 27	22.2	18.0	80	9	CL302865	CL302865 P015B08 G
C 28	22.2	18.0	88	9	CG869072	CG869072 AB0246 Sa
C 29	22.2	18.0	93	8	CC178569	CC178569 NPX449 Ba
C 30	22.2	18.0	102	4	BI046042	BI046042 MR3-FN020
31	22.2	18.0	109	1	AW000596	AW000596 614056B09
32	22.2	18.0	114	2	BE165383	BE165383 QV1-HT047
C 33	22.2	18.0	114	9	CG869066	CG869066 AB0239 Sa
C 34	22.2	18.0	119	6	BY641401	BY641401 BY641401
C 35	22	17.9	66	8	AZ656200	AZ656200 1M0531A10
36	22	17.9	110	4	BM484534	BM484534 538432 MA
37	22	17.9	111	4	BG274104	BG274104 WHE2231 F
38	22	17.9	120	1	AI617421	AI617421 zehnl542.
C 39	22	17.9	123	2	AW137016	AW137016 UI-H-B11-
C 40	22	17.9	123	8	BZ705049	BZ705049 PUCBW34TD
41	21.8	17.7	91	1	AA639724	AA639724 nq93d05.s
42	21.8	17.7	113	9	CG477283	CG477283 OST7129 M
43	21.8	17.7	113	9	CG484766	CG484766 OST18836
44	21.8	17.7	116	2	BF019263	BF019263 uy03b04.Y
C 45	21.8	17.7	117	9	CC713422	CC713422 OGUDH54TH

ALIGNMENTS

RESULT 1
AA430669/c
LOCUS zw26d03.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
DEFINITION IMAGE:770405 3', mRNA sequence.
ACCESSION AA430669
VERSION AA430669.1 GI:2111242
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 97)
AUTHORS Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyT not found
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 65.

FEATURES
source
1. .97
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5980279"
/db_xref="taxon:9606"
/clone="IMAGE:770405"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares ovary tumor NbHOT"
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGCGCGCGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Ronaldo. "

ORIGIN

	Query Match	21.0%;	Score 25.8; DB 1;	Length 97;	
	Best Local Similarity	56.5%;	Pred. No. 4.7e+03;		
	Matches 48; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;
QY	39	CCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTCGCCCGGCATGTGCCTGCAGT	98		
Dd	97	CCCCAGCCGGCGCCTGCCCTCTTTCTCCCGGCCACTGCCAACCTCCAGTGCCCGCCGAGA	38		
QY	99	GTGGGTGCTGCTCAGCCGGGACCCTC	123		
Dd	37	ACCTCCGCTGCCTCGGCGCGACCCCC	113		

RESULT 2
 BF171548/c
 LOCUS
 DEFINITION
 117 bp mRNA linear EST 23-MAR-2001
 PCL2912 Myeloma (PCL) cDNA library Homo sapiens cDNA, mRNA
 sequence.
 BF171548
 BF171548.1 GI:13437762
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 117)
 REFERENCE
 AUTHORS
 Claudio, J.O., Masih-Khan, E., Tang, H., Goncalves, J., Voralia, M.,
 Li, Z.H., Nadeem, V., Cukerman, E., Francisco-Pabalan, O., Liew, C.C.,
 Woodgett, J.R. and Stewart, A.K.
 TITLE
 A molecular compendium of genes expressed in multiple myeloma
 JOURNAL
 Blood 100 (6), 2175-2186 (2002)
 MEDLINE
 22188429
 PUBMED
 12200383
 COMMENT
 Contact: A. Keith Stewart, M.D.

FEATURES
SOURCE

ORIGIN

Query Match	21.0%;	Score 25.8;	DB 2;	Length 117;
Best Local Similarity	56.5%;	Pred. NO. 4.8e+03;		
Matches 48;	Conservative 0;	Mismatches 37;	Indels 0;	Gaps 0;
QY	11	GTTCCTTCAGACTGAGGACAGCTGGGTCCACAGCCTCACCTGACACAGGGCTAGACCCCC	70	
Db	101	GTCTCGTCTAACACCAAGCTAGAAGGGGCGGGGCCCTCCCTCAAGCTGGGGCTCGGCCTCC	42	
QY	71	TCACAGTGGCGGCCCATGTGCCTGC	95	
Db	41	TCCCCGCGCGGCGAGCTGGTGCCTCC	17	

RESULT 3	
BY074656/c	
LOCUS	118 bp mRNA linear EST 06-DEC-2002
DEFINITION	BY074656 RIKEN full-length enriched, 16 days embryo heart Mus musculus cDNA clone I920096L06 5', mRNA sequence.
ACCESSION	BY074656
VERSION	BY074656.1 GI:26176407
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 118)
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Db 47 TAGGTGCTTTTGGCGGAGGACAGGTTCAGGCCCTGGATACTCTTACACAGCGCAATAAG 106

QY 68 CCC 70

Db 107 ACC 109

RESULT 6

AZ883677/c

LOCUS

DEFINITION

RPCI-23-187D12.TJ RPCI-23 Mus musculus genomic clone linear GSS 05-MAR-2001

RPCI-23-187D12, genomic survey sequence.

ACCESSION

AZ883677

VERSION

AZ883677.1 GI:13202622

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AZ883677

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 112)

Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.

TITLE

Mouse BAC End Sequences from Library RPCI-23

JOURNAL

Unpublished (1999)

COMMENT

Other GSSs: RPCI-23-187D12.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end page: http://ww.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 187 row: D column: 12

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..112

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-187D12"

/sex="Female"

/lab_host="DH10B"

/clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 19.8%; Score 24.4; DB 8; Length 112;

Best Local Similarity 68.0%; Pred. No. 1.2e+04;

Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 19 CAGACTGAGCAGCTGGGTCCAGCCTCACCTGACACAGGGCTAGACCC 68

Db 110 CAGACTGAGAAAGATCTTTCCAATCTTCATCCAAAGAGGGCTCTATCC 61

RESULT 7

LBAF042F12

LOCUS

LBAF042F12 115 bp DNA linear GSS 23-JUN-2003

DEFINITION

Leishmania braziliensis GSS, clone LBAF042F12, genomic survey sequence.

ACCESSION

BX545265

VERSION

BX545265.1 GI:32170063

KEYWORDS

GSS; genomic survey sequence.

SOURCE

Leishmania braziliensis

ORGANISM

Leishmania braziliensis

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania; Leishmania braziliensis species complex.

REFERENCE

1

AUTHORS

Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.

TITLE

GSS analysis of the Leishmania braziliensis genome

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 115)

AUTHORS

Cruz,A.K.

TITLE

Direct Submission

JOURNAL

Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo, Department of Molecular and Cell Biology, FMRP, Avenida Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL

Clone requests: akcruz@mrp.usp.br.

Location/Qualifiers

1..115

/organism="Leishmania braziliensis"

/mol_type="genomic DNA"

/strain="MHOM/BR/75/M2904"

/db_xref="taxon:5660"

/clone="LBAF042F12"

ORIGIN

Query Match 19.7%; Score 24.2; DB 9; Length 115;

Best Local Similarity 71.1%; Pred. No. 1.4e+04;

Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 60 GCTAGACCCCTCACAGTGCCTGCGCCCATGTGCCTGCAGTGTGGGT 104

Db 71 GGTAAAGCCCTTCAGCCTGCGCAGACACCTGCGTGTAGTGTGGGT 115

RESULT 8

AI262761

LOCUS

AI262761

DEFINITION

AI262761.x1 NCI CGAP Kid3 Homo sapiens cDNA clone IMAGE:1868206 3, similar to SW:PRP4_HUMAN P10163 SALIVARY PROLINE-RICH PROTEIN PO PRECURSOR ;contains PTRS.b3 MSR1 repetitive element ;, mRNA sequence.

ACCESSION

AI262761

VERSION

AI262761.1 GI:3870964

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 52)

REFERENCE

1

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40UP from Gibco

High quality sequence stop: 1.

Location/Qualifiers

1..52

FEATURES

source

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 20:17:18 ; Search time 269 Seconds
(without alignments)
2400.293 Million cell updates/sec

Title: US-10-057-813-13_COPY_253_375
Perfect score: 123
Sequence: 1 accaagctacgtttccttca.....tgctgctcagccgggacccc 123

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 4444388

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 123

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25.8	21.0	97	6	ABL81700 Human ova
2	25.2	20.5	121	10	ADF86506 Single nu
3	24.4	19.8	100	8	ACD79058 E. coli K
4	23.2	18.9	118	10	ADC40305 Human sec
C 5	23	18.7	89	3	AAC11438 Human sec
6	23	18.7	121	4	ABA77492
C 7	23	18.7	121	4	ABA77491
C 8	22.8	18.5	121	10	ADF87025
C 9	22.6	18.4	121	4	AAK70400
C 10	22.6	18.4	121	4	AAK70397
C 11	22.2	18.0	121	10	ADF87024
12	21.8	17.7	100	8	ACD76144
13	21.8	17.7	111	12	ACH89320 Human gen
14	21.6	17.6	120	2	AAX12744 Human bia
C 15	21.4	17.4	119	3	AAC11139 Human sec
C 16	21.4	17.4	121	12	ADL80264 Human rec
17	21.2	17.2	91	6	ABN63065 Human can
18	21.2	17.2	100	8	ACD79057
C 19	21.2	17.2	121	4	ABA77363
20	21.2	17.2	121	4	ABA77364
C 21	21	17.1	58	6	ABL32014 ApoE2150-

C 22	21	17.1	65	6	ABN54356	Abn54356 Mouse spl
23	21	17.1	70	12	ADP05517	Adp05517 Novel mic
C 24	21	17.1	100	8	ACD70591	Acd70591 E. coli K
25	21	17.1	121	12	ADK91981	Adk91981 Polynucle
C 26	20.8	16.9	54	10	ADC35401	Adc35401 PCR prime
27	20.8	16.9	60	6	ABN59212	Abn59212 Human spl
28	20.8	16.9	60	6	ABN59043	Abn59043 Human spl
C 29	20.8	16.9	100	8	ACD72441	Acd72441 E. coli K
C 30	20.8	16.9	106	12	ACH85499	Ach85499 Human gen
C 31	20.8	16.9	112	4	AAH34256	Aah34256 Human col
32	20.8	16.9	121	12	ADK91466	Adk91466 Polynucle
33	20.8	16.9	121	12	ADL80566	Adl80566 Human rec
34	20.6	16.7	65	6	ABN57458	Abn57458 Mouse spl
35	20.6	16.7	77	3	AAC20910	Aac20910 Human sec
C 36	20.6	16.7	80	12	ADM96137	Adm96137 Rat antis
C 37	20.6	16.7	112	4	AAI24307	Aai24307 Probe #14
C 38	20.6	16.7	112	4	ABA69447	Aba69447 Human foe
C 39	20.6	16.7	112	4	AAK43526	Aak43526 Human bon
C 40	20.6	16.7	112	5	AAI09857	Aai09857 Probe #98
C 41	20.6	16.7	112	6	ABS17656	Abas17656 Human gen
42	20.6	16.7	121	10	ADH92474	Adh92474 Human gen
C 43	20.6	16.7	121	10	ADH92553	Adh92553 Human gen
C 44	20.4	16.6	57	12	ADN35911	Adn35911 Human Reg
C 45	20.4	16.6	65	6	ABN53195	Abn53195 Mouse spl

ALIGNMENTS

RESULT 1
ABL81700/c
ID ABL81700 standard; cDNA; 97 BP.
XX
AC ABL81700;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:4678.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WQ200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
PS Claim 1; SEQ ID NO 4678; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological

sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (III) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques

SQ Sequence 97 BP; 15 A; 18 C; 53 G; 11 T; 0 U; 0 Other;
 Query Match 21.0%; Score 25.8; DB 6; Length 97;
 Best Local Similarity 56.5%; Pred. No. 1.9e+02;
 Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

RESULT 2
ADFF86506
ID ADF86506 standard; DNA; 121 BP.
XX
AC ADF86506;
XX
DT 26-FEB-2004 (first entry)
XX
DE Single nucleotide polymorphism detection human gene, SEQ ID No 89.
XX
KW human; single nucleotide polymorphism; microarray; side effect; gene; ds.

The invention relates to a novel polynucleotide isolated and purified from a human gene having any one of 935 fully defined sequences as given in specification, or a sequence having a base substitution. The invention further relates to: an oligonucleotide containing single nucleotide polymorphisms; a PCR primer set chosen from the combination of two DNA fragments from any one of 1220 fully defined sequences as given in specification; a labelling probe containing the SNP containing oligo; and a microarray equipped with the SNP containing oligo. The isolated human gene of the invention is useful for detecting the single nucleotide polymorphisms in human gene. The isolated human gene is also useful for diagnosis of disease and determination of side effect to a medical agent. The isolated human gene is also effective in detecting single nucleotide polymorphisms in a human gene. This polynucleotide sequence represents one of the 935 isolated polynucleotides from a human gene of the invention.

```
XX  
SQ      Sequence 121 BP; 18 A; 47 C; 31 G; 24 T; 0 U; 1 Other;  
  
Query Match          20.5%;   Score 25.2;    DB 10; Length 121;  
Best Local Similarity 54.5%;   Pred. No. 3.le+02;  
Matches 48; Conservative 1; Mismatches 39; Indels 0; Gaps 0;
```

```
QY      33 CTGGGTCCCAGCCTCACCTGACACAGGGGTAGACCCCTTCACAGTGC CGCGCATGTGCC 92  
||||| ||||| ||||| ||||| : |||| | ||| ||||  
Db       29 CTGGGTCCCCAATCTCCTGGACGGGCCCTCRCA CCCAGCCCTTTTGCA 88  
  
QY      93 TGCA GTGTGGTGTCTGCTC AGCCGGGAC 120  
- - - - - |||| | - - - - -  
Db      89 GGTC TTGTCCATGGAGA GAAGCTGGAC 116
```

RESULT 3
ACD79058
ID ACD79058 standard; DNA; 100 BP.
XX
AC ACD79058;
XX
DT 19-SEP-2003 (first entry)
XX
DE E. coli K12 MG1655 biochip probe SEQ ID 10334.
XX
KW Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX Escherichia coli.
OS
PN EP1260592-A1.
XX
PD 27-NOV-2002.
XX
PF 17-MAY-2001; 2001EP-00112179.
XX
PR 17-MAY-2001; 2001EP-00112179.
XX
PA (MWGB-) MWG-BIOTECH AG.
XX
PI Donner H, Drescher B, Huber A, Weber J;
XX WPI; 2003-241155/24.
DR

This invention describes a novel biochip comprising probe spots, each containing many identical probes. The probes are nucleotide sequences of 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at least one includes a segment of at least 20 bases identical with, or complementary to, a segment of an open reading frame (orf) of *Escherichia coli* K12. The biochip is used for specific detection of gene expression in K12 and for determining the gene expression pattern, e.g. for diagnostic determination of which *E. coli* strains are present in the gut, and to determine the effects of e.g. growth media on gene expression. The biochip provides as comprehensive as possible detection of the K12 genome, with simultaneous analysis of many different genes with a single device, and comparison of gene expression between K12 and its mutants or other *E. coli* strains in a single experiment. Apart from qualitative and quantitative information about gene expression, it also allows measurements of population densities for the various strains. The use of synthetic oligonucleotides for preparation of probes allows free variation in probe length and ensures high purity (and thus selectivity, reactivity and reproducibility); also synthetic probes are generally shorter than probes prepared by polymerase chain reaction. ACD68731 to ACD81540 represent oligonucleotide probes used with the biochip described in the invention

Sequence 100 BP; 23 A; 37 C; 27 G; 13 T; 0 U; 0 Other;

KW UDP-glucuronosyltransferase; amyloid precursor protein; presenilin-1;
KW Alzheimer's disease; cytostatic; antisickling; antianaemic; haemostatic;
XX antilipemic; ss.
OS Homo sapiens.
XX
PN WO200173002-A2.
PD 04-OCT-2001.
XX
XX
PF 27-MAR-2001; 2001WO-US009761.
XX
XX 27-MAR-2000; 2000US-0192176P.
PR 27-MAR-2000; 2000US-0192179P.
PR 01-JUN-2000; 2000US-0208538P.
PR 30-OCT-2000; 2000US-0244989P.
XX
PA (UYDE) UNIV DELAWARE.
XX
XX Kmiec EB, Gamper HB, Rice MC;
PI
XX
DR WPI; 2001-639230/73.
XX
XX Oligonucleotide for targeted alterations of genetic sequences and for
PT treating cystic fibrosis, comprises at least one mismatch and chemical
PT modification.
XX
PS Claim 7; Page 63; 294pp; English.
XX
XX The present invention provides single-stranded oligonucleotides which can
CC be used for the targeted alteration of genomic sequences, where the
CC oligonucleotide has at least one mismatch compared with the genomic
CC sequence to be altered. In particular, these sequences are directed at
CC the following genes: adenosine deaminase, p53, beta-globin,
CC retinoblastoma, BRCA1, BRCA2, CFTR, cyclin-dependent kinase inhibitor 2A
CC (CDKN2A), APC, Factor V, Factor VIII, Factor IX, haemoglobin alpha locus
CC 1'(HBA1), haemoglobin alpha locus 2 (HBA2), MLH1, MSH2, MSH6,
CC apolipoprotein E (APOE), LDL receptor (LDLR), UDP-glucuronosyltransferase
CC (UGT1), amyloid precursor protein (APC), presenilin-1 (PSEN1) and
CC presenilin-2 (PSEN2). These can be used in the gene therapy of diseases
CC such as cancer, adenosine deaminase deficiency, cystic fibrosis,
CC haemophilia, hypercholesterolaemia, thalassaemia, sickle cell anaemia,
CC Alzheimer's disease, melanoma, adenomatous polyposis of the colon and
CC various syndromes. The present sequence is one of the gene correcting
CC oligonucleotides of the invention
XX
SQ Sequence 121 BP; 17 A; 43 C; 31 G; 30 T; 0 U; 0 Other;
Query Match 18.7%; Score 23; DB 4; Length 121;
Best Local Similarity 54.0%; Pred. No. 1.5e+03;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCGCTCACCTGACACAGGGCTAGACCCC 69
Db 4 CGCTTAGTGCTCCCTGGGGCAGCTCGTGGTGAGGCTCCCTTTCTTGGGAGATTCTCT 63
QY 70 CTCACAGTGGCGCCGCTGTGCCTGCA 96
Db 64 TCCTCTGTGCGCGGTCTCTCCAGGA 90
RESULT 7
ABA77491/c
ID ABA77491 standard; DNA; 121 BP.
XX
AC ABA77491;
XX
DT 24-JAN-2002 (first entry)
XX
DE p53 mutation correcting oligonucleotide SEQ ID NO: 337.
XX
KW Human; gene therapy; adenosine deaminase deficiency; p53; beta-globin;
KW retinoblastoma; BRCA1; BRCA2; CFTR; cystic fibrosis; cancer; Factor V;

KW cyclin-dependent kinase inhibitor 2A; CDKN2A; melanoma; APC; HBA1; HBA2;
KW adenomatous polyposis of the colon; Factor VII; Factor IX; thrombosis;
KW haemophilia; alpha thalassaemia; haemoglobin alpha locus 1; MLH1; APOE;
KW mismatch repair; MSH2; MSH6; hyperlipidaemia; apolipoprotein E; LDLR;
KW familial hypercholesterolaemia; UGT1; syndrome; APP; PSEN1; antisense;
KW UDP-glucuronosyltransferase; amyloid precursor protein; presenilin-1;
KW Alzheimer's disease; cytostatic; antisickling; antianaemic; haemostatic;
XX antilipemic; ss.
XX
OS Homo sapiens.
XX
PN WO200173002-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US009761.
XX
XX 27-MAR-2000; 2000US-0192176P.
PR 27-MAR-2000; 2000US-0192179P.
PR 01-JUN-2000; 2000US-0208538P.
PR 30-OCT-2000; 2000US-0244989P.
XX
PA (UYDE) UNIV DELAWARE.
XX
XX Kmiec EB, Gamper HB, Rice MC;
PI
XX
DR WPI; 2001-639230/73.
XX
XX Oligonucleotide for targeted alterations of genetic sequences and for
PT treating cystic fibrosis, comprises at least one mismatch and chemical
PT modification.
XX
PS Claim 7; Page 63; 294pp; English.
XX
XX The present invention provides single-stranded oligonucleotides which can
CC be used for the targeted alteration of genomic sequences, where the
CC oligonucleotide has at least one mismatch compared with the genomic
CC sequence to be altered. In particular, these sequences are directed at
CC the following genes: adenosine deaminase, p53, beta-globin,
CC retinoblastoma, BRCA1, BRCA2, CFTR, cyclin-dependent kinase inhibitor 2A
CC (CDKN2A), APC, Factor V, Factor VIII, Factor IX, haemoglobin alpha locus
CC 1 (HBA1), haemoglobin alpha locus 2 (HBA2), MLH1, MSH2, MSH6,
CC apolipoprotein E (APOE), LDL receptor (LDLR), UDP-glucuronosyltransferase
CC (UGT1), amyloid precursor protein (APC), presenilin-1 (PSEN1) and
CC presenilin-2 (PSEN2). These can be used in the gene therapy of diseases
CC such as cancer, adenosine deaminase deficiency, cystic fibrosis,
CC haemophilia, hypercholesterolaemia, thalassaemia, sickle cell anaemia,
CC Alzheimer's disease, melanoma, adenomatous polyposis of the colon and
CC various syndromes. The present sequence is one of the gene correcting
CC oligonucleotides of the invention
XX
SQ Sequence 121 BP; 30 A; 31 C; 43 G; 17 T; 0 U; 0 Other;
Query Match 18.7%; Score 23; DB 4; Length 121;
Best Local Similarity 54.0%; Pred. No. 1.5e+03;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCCGCTCACCTGACACAGGGCTAGACCCC 69
Db 118 CGCTTAGTGCTCCCTGGGGCAGCTCGTGGTGAGGCTCCCTTTCTTGGGAGATTCTCT 59
QY 70 CTCACAGTGGCGCCGCTGTGCCTGCA 96
Db 58 TCCTCTGTGCGCGGTCTCTCCAGGA 32
RESULT 8
ADF87025/c
ID ADF87025 standard; DNA; 121 BP.
XX
AC ADF87025;
XX
DT 26-FEB-2004 (first entry)

XX Single nucleotide polymorphism detection human gene, SEQ ID No 608.
DE human; single nucleotide polymorphism; microarray; side effect; gene; ds.
XX Homo sapiens.
XX JP2003235571-A.
XX 26-AUG-2003.
XX 12-FEB-2002; 2002JP-00034717.
XX 12-FEB-2002; 2002JP-00034717.
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX WPI; 2003-820454/77.
XX Novel polymucleotide useful for detecting single nucleotide polymorphisms in human gene.
XX Claim 1; SEQ ID NO 608; 704pp; Japanese.
XX The invention relates to a novel polynucleotide isolated and purified from a human gene having any one of 935 fully defined sequences as given in specification, or a sequence having a base substitution. The invention further relates to: an oligonucleotide containing single nucleotide polymorphisms; a PCR primer set chosen from the combination of two DNA fragments from any one of 1220 fully defined sequences as given in specification; a labelling probe containing the SNP containing oligo; and a microarray equipped with the SNP containing oligo. The isolated human gene of the invention is useful for detecting the single nucleotide polymorphisms in human gene. The isolated human gene is also useful for diagnosis of disease and determination of side effect to a medical agent. The isolated human gene is also effective in detecting single nucleotide polymorphisms in a human gene. This polynucleotide sequence represents one of the 935 isolated polynucleotides from a human gene of the invention.

XX SQ Sequence 121 BP; 23 A; 40 C; 43 G; 14 T; 0 U; 1 Other;
Query Match 18.5%; Score 22.8; DB 10; Length 121;
Best Local Similarity 75.0%; Pred.No.1.8e+03;
Matches 27; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 64 GACCCCTCAGTGGCGCCCATGTGCTGCAGTG 99
:|||||
Db 61 KACCGCTCAGATTGCTCCACCAAGGTGCTGTGGG 26

RESULT 9
AAK70400/c
ID AAK70400 standard; DNA; 121 BP.
XX AAK70400;
AC
XX 06-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25212.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
XX 09-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001354.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR	02-OCT-2000;	2000US-02370308P
PR	02-OCT-2000;	2000US-02370339P
PR	02-OCT-2000;	2000US-02370400P
PR	13-OCT-2000;	2000US-02393355P
PR	13-OCT-2000;	2000US-0239337P
PR	20-OCT-2000;	2000US-0240960P
PR	20-OCT-2000;	2000US-0241221P
PR	20-OCT-2000;	2000US-0241785P
PR	20-OCT-2000;	2000US-0241786P
PR	20-OCT-2000;	2000US-0241787P
PR	20-OCT-2000;	2000US-0241808P
PR	20-OCT-2000;	2000US-0241809P
PR	20-OCT-2000;	2000US-0241826P
PR	01-NOV-2000;	2000US-0244617P
PR	08-NOV-2000;	2000US-0246474P
PR	08-NOV-2000;	2000US-0246475P
PR	08-NOV-2000;	2000US-0246476P
PR	08-NOV-2000;	2000US-0246477P
PR	08-NOV-2000;	2000US-0246478P
PR	08-NOV-2000;	2000US-0246523P
PR	08-NOV-2000;	2000US-0246524P
PR	08-NOV-2000;	2000US-0246525P
PR	08-NOV-2000;	2000US-0246526P
PR	08-NOV-2000;	2000US-0246527P
PR	08-NOV-2000;	2000US-0246528P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-0246609P
PR	08-NOV-2000;	2000US-0246610P
PR	08-NOV-2000;	2000US-0246611P
PR	08-NOV-2000;	2000US-0246613P
PR	17-NOV-2000;	2000US-0249207P
PR	17-NOV-2000;	2000US-0249208P
PR	17-NOV-2000;	2000US-0249209P
PR	17-NOV-2000;	2000US-0249210P
PR	17-NOV-2000;	2000US-0249211P
PR	17-NOV-2000;	2000US-0249212P
PR	17-NOV-2000;	2000US-0249213P
PR	17-NOV-2000;	2000US-0249214P
PR	17-NOV-2000;	2000US-0249215P
PR	17-NOV-2000;	2000US-0249216P
PR	17-NOV-2000;	2000US-0249217P
PR	17-NOV-2000;	2000US-0249218P
PR	17-NOV-2000;	2000US-0249244P
PR	17-NOV-2000;	2000US-0249245P
PR	17-NOV-2000;	2000US-0249264P
PR	17-NOV-2000;	2000US-0249265P
PR	17-NOV-2000;	2000US-0249297P
PR	17-NOV-2000;	2000US-0249299P
PR	17-NOV-2000;	2000US-0249300P
PR	01-DEC-2000;	2000US-0250160P
PR	01-DEC-2000;	2000US-0250391P
PR	05-DEC-2000;	2000US-0251030P
PR	05-DEC-2000;	2000US-0251988P
PR	05-DEC-2000;	2000US-0256719P
PR	06-DEC-2000;	2000US-0251479P
PR	08-DEC-2000;	2000US-0251856P
PR	08-DEC-2000;	2000US-0251868P
PR	08-DEC-2000;	2000US-0251869P
PR	08-DEC-2000;	2000US-0251989P
PR	11-DEC-2000;	2000US-0251990P
PR	05-JAN-2001;	2001US-0259678P

XX

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention

Sequence 121 BP; 34 A; 19 C; 56 G; 12 T; 0 U; 0 Other;

PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 25209; 3071pp + Sequence Listing; English.
XX

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169

CC represent sequences used in the exemplification of the present invention

XX

SQ Sequence 121 BP; 34 A; 19 C; 56 G; 12 T; 0 U; 0 Other;

Query Match 18.4%; Score 22.6; DB 4; Length 121;

Best Local Similarity 60.7%; Pred. No. 2.1e+03;

Matches 37; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 14 TCCTTCAGACTGAGGACAGCTGGGTCCCAGCCTCACCTGACAGGGCTAGACCCCTCA 73

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

65 TGCCTCAGACTGACACTAGCTTCTTCACTGACCTCTCTGCTGCAGGGCTGTCCCTCTGA 6

QY 74 C 74

Db |

5 C 5

RESULT 11

ADF87024/c

ID ADF87024 standard; DNA; 121 BP.

XX

AC ADF87024;

XX

DT 26-FEB-2004 (first entry)

XX

DE Single nucleotide polymorphism detection human gene, SEQ ID No 607.

XX

KW human; single nucleotide polymorphism; microarray; side effect; gene; ds.

XX

OS Homo sapiens.

XX

PN JP2003235571-A.

XX

PD 26-AUG-2003.

XX

PF 12-FEB-2002; 2002JP-00034717.

XX

PR 12-FEB-2002; 2002JP-00034717.

XX

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX

DR WPI; 2003-820454/77.

XX

PT Novel polynucleotide useful for detecting single nucleotide polymorphisms

PT in human gene.

XX

PS Claim 1; SEQ ID NO 607; 704pp; Japanese.

XX

CC The invention relates to a novel polynucleotide isolated and purified

CC from a human gene having any one of 935 fully defined sequences as given

CC in specification, or a sequence having a base substitution. The invention

CC further relates to: an oligonucleotide containing single nucleotide

CC polymorphisms; a PCR primer set chosen from the combination of two DNA

CC fragments from any one of 1220 fully defined sequences as given in

CC specification; a labelling probe containing the SNP containing oligo; and

CC a microarray equipped with the SNP containing oligo. The isolated human

CC gene of the invention is useful for detecting the single nucleotide

CC polymorphisms in human gene. The isolated human gene is also useful for

CC diagnosis of disease and determination of side effect to a medical agent.

CC The isolated human gene is also effective in detecting single nucleotide

CC polymorphisms in a human gene. This polynucleotide sequence represents

CC one of the 935 isolated polynucleotides from a human gene of the

CC invention.

XX

SQ Sequence 121 BP; 24 A; 42 C; 40 G; 14 T; 0 U; 1 Other;

Query Match 18.0%; Score 22.2; DB 10; Length 121;

Best Local Similarity 77.1%; Pred. No. 2.8e+03;

Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 65 ACCCCCTCACAGTGGCGCCGATGTGCCTGCAGTG 99

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

57 ACCGCTCACATTGCTCCACAGGTGCCTGTGGGG 23

RESULT 13
ACH89320
ID ACH89320 standard; DNA; 111 BP.
XX
AC ACH89320;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #22515.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 22515; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

```

XX
SQ   Sequence 111 BP; 18 A; 35 C; 30 G; 28 T; 0 U; 0 Other;

Query Match      17.7%;   Score 21.8;   DB 12;   Length 111;
Best Local Similarity 70.7%;   Pred. No. 3.6e+03;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY   65 ACCCCCTCACAGTGGCGCCCATGTGCGCTGCAGTGTGGGTG 105
    |||||
Db   19 AGCCCTCACAGAGCTTGCTCCTGTGTGTCAGCAGAGTGGCTG 59

RESULT 14
AAX12744
ID   AAX12744 standard; DNA; 120 BP.
XX
AC   AAX12744;
XX
DT   30-MAR-1999 (first entry)
XX
DE   Human biallelic polymorphic DNA fragment stSG3935.
XX
KW   Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW   detection; phenotypic typing; characteristic; infection; hereditary;
KW   autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW   treatment; marker; ss.
XX
OS   Homo sapiens.
XX
PN   WO9820165-A2.
XX
PD   14-MAY-1998.
XX
PF   05-NOV-1997; 97WO-US020313.
XX
PR   06-NOV-1996; 96US-0030455P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lander ES, Wang D, Hudson T;
XX
DR   WPI; 1998-286974/25.
XX
PT   New isolated nucleic acid segments from the human genome - used for
PT   determining polymorphic forms for use in e.g. forensics, paternity
PT   testing or phenotypic typing for disease.
XX
PS   Claim 1; Page 286; 310pp; English.
XX
CC   AAX10269-X12937 are human DNA fragments which contain biallelic
CC   polymorphic markers which have been isolated using the primers
CC   represented in AAX09121-X10268. The base occupying the polymorphic site
CC   is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC   can be used in methods for determining polymorphic forms in an individual
CC   for use in e.g. forensics, paternity testing or for phenotypic typing for
CC   diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC   syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC   familial hypercholesterolemia, polycystic kidney disease, hereditary
CC   spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC   haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC   syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC   autoimmune diseases, inflammation, cancer, diseases of the nervous
CC   system, infection by pathogenic microorganisms, and characteristics such
CC   as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC   endurance, fertility, and susceptibility or receptivity to particular
CC   drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC   segments can also be used to produce medicaments for the treatment or
CC   prophylaxis of such diseases
XX
SQ   Sequence 120 BP; 25 A; 37 C; 30 G; 27 T; 0 U; 1 Other;

Query Match      17.6%;   Score 21.6;   DB 2;   Length 120;
Best Local Similarity 68.2%;   Pred. No. 4.2e+03;

```

Matches 30; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 11 GTTTCCTTCAGACTGAGGACAGCTGGGTCCCGCTCACCTGAC 54
 ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 75 GTTCCACAGGCTGAGTACTATGGGGTCACAACTTCCTGGAC 118

RESULT 15
AAC11139/c
ID AAC11139 standard; cDNA; 119 BP.
XX
AC AAC11139;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 15214.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX

OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX

PS Claim 1; SEQ ID NO 15214; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX

SQ Sequence 119 BP; 19 A; 44 C; 32 G; 24 T; 0 U; 0 Other;

Query Match 17.4%; Score 21.4; DB 3; Length 119;
Best Local Similarity 61.8%; Pred. No. 4.9e+03;
Matches 34; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 69 CCTCAGTGGCGCGCATGTGCCTGCAGTGTGGTGCTGTCTAGCCGGACCCC 123
 ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 76 CATGAACGTGGCAGGAAGCGGTGCAATGTGGGCTGTTCTGTCATGGAGCGC 22

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 21:43:24 ; Search time 307 Seconds
(without alignments)
2302.099 Million cell updates/sec

Title: US-10-057-813-13_COPY_253_375
Perfect score: 123
Sequence: 1 accaagctacgtttcctca.....tgctgtcagccgggacccc 123

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 3219360

Minimum DB seq length: 0
Maximum DB seq length: 123

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.8	21.0	97	US-09-867-701-4678	Sequence 4678, Ap
C 2	23.2	18.9	101	US-10-674-124A-21157	Sequence 21157, A
C 3	23	18.7	121	US-09-818-875-337	Sequence 337, App
C 4	23	18.7	121	US-09-818-875-338	Sequence 338, App
C 5	23	18.7	121	US-10-209-787-337	Sequence 337, App
C 6	23	18.7	121	US-10-209-787-338	Sequence 338, App
C 7	23	18.7	121	US-10-261-185-337	Sequence 337, App
C 8	23	18.7	121	US-10-261-185-338	Sequence 338, App
C 9	23	18.7	121	US-10-681-074-337	Sequence 337, App
C 10	23	18.7	121	US-10-681-074-338	Sequence 338, App
C 11	22.2	18.0	117	US-10-242-535A-10710	Sequence 10710, A
C 12	22.2	18.0	117	US-10-085-783A-10710	Sequence 10710, A

C 13	21.8	17.7	111	15	US-10-029-386-22515	Sequence 22515, A
C 14	21.2	17.2	121	10	US-09-818-875-209	Sequence 209, App
C 15	21.2	17.2	121	10	US-09-818-875-210	Sequence 210, App
C 16	21.2	17.2	121	15	US-10-209-787-209	Sequence 209, App
C 17	21.2	17.2	121	15	US-10-209-787-210	Sequence 210, App
C 18	21.2	17.2	121	16	US-10-261-185-209	Sequence 209, App
C 19	21.2	17.2	121	16	US-10-261-185-210	Sequence 210, App
C 20	21.2	17.2	121	17	US-10-681-074-209	Sequence 209, App
C 21	21.2	17.2	121	17	US-10-681-074-210	Sequence 210, App
C 22	21	17.1	65	10	US-09-908-975-27104	Sequence 27104, A
C 23	20.8	16.9	60	10	US-09-908-975-31791	Sequence 31791, A
C 24	20.8	16.9	60	10	US-09-908-975-31960	Sequence 31960, A
C 25	20.8	16.9	106	15	US-10-029-386-18694	Sequence 18694, A
C 26	20.8	16.9	112	15	US-10-106-698-1348	Sequence 1348, Ap
C 27	20.6	16.7	65	10	US-09-908-975-30206	Sequence 30206, A
C 28	20.6	16.7	80	16	US-10-384-245-1030	Sequence 1030, Ap
C 29	20.6	16.7	112	9	US-09-864-761-31659	Sequence 31659, A
C 30	20.6	16.7	114	18	US-10-674-124A-24132	Sequence 24132, A
C 31	20.4	16.6	65	10	US-09-908-975-25943	Sequence 25943, A
C 32	20.4	16.6	65	10	US-09-908-975-27954	Sequence 27954, A
C 33	20.4	16.6	101	18	US-10-674-124A-21552	Sequence 21552, A
C 34	20.4	16.6	119	17	US-10-437-963-88002	Sequence 88002, A
C 35	20.2	16.4	60	10	US-09-908-975-11338	Sequence 11338, A
C 36	20.2	16.4	96	15	US-10-029-386-27083	Sequence 27083, A
C 37	20.2	16.4	122	18	US-10-425-115-163501	Sequence 163501, A
C 38	20.2	16.4	123	15	US-10-029-386-19292	Sequence 19292, A
C 39	20	16.3	96	9	US-09-915-873-3	Sequence 3, Appli
C 40	20	16.3	96	18	US-10-794-615-3	Sequence 32748, A
C 41	20	16.3	105	9	US-09-864-761-32748	Sequence 382, App
C 42	19.8	16.1	69	11	US-09-837-306-382	Sequence 383, App
C 43	19.8	16.1	69	11	US-09-837-306-383	Sequence 298, App
C 44	19.8	16.1	69	15	US-10-045-674-298	Sequence 299, App
C 45	19.8	16.1	69	15	US-10-045-674-299	

ALIGNMENTS

RESULT 1

US-09-867-701-4678/c

; Sequence 4678, Application US/09867701

; Patent No. US20020132237A1

; GENERAL INFORMATION:

; APPLICANT: Aglate, Paul A.

; APPLICANT: Jones, Robert

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

; FILE REFERENCE: 210121.497

; CURRENT APPLICATION NUMBER: US/09/867,701

; CURRENT FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 10912

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4678

; LENGTH: 97

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-867-701-4678

Query Match	21.0%;	Score 25.8;	DB 9;	Length 97;
Best Local Similarity	56.5%;	Pred. No. 76;		
Matches	48;	Conservative	0;	Mismatches 37; Indels 0; Gaps 0;
Qy	39	CCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGCCTGCAGT	98	
Dd	97	CCCCCGCGCCTGCCTCTCTTCTCCCGCCACTGCCAACCTCCAGTCGCCCGCAG	38	
Qy	99	GTGGGTGCTGCTCAGCGGGACCCC	123	
Dd	37	ACCTCCGCTGCTCGCGCGACCCC	13	
RESULT 2				


```
US-10-674-124A-21157/c
; Sequence 21157, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 21157
; LENGTH: 101
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: DID22N_0010585
; OTHER INFORMATION: Located on chromosome 14
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 104184465
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 76372
US-10-674-124A-21157

Query Match      18.9%; Score 23.2; DB 18; Length 101;
Best Local Similarity 56.6%; Pred. No. 5.5e+02;
Matches 43; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 12 TTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCAGCTGACACAGGGCTAGACCCCT 71
Db 101 TCTCCTCCTAAATGAGCAGTCCCTCACCCTGCTTCCCTTCACACACACACACACA 42

QY 72 CACAGTGCGCCGCAT 87
Db 41 CACTGTGTGTCCAT 26

RESULT 3
US-09-818-875-337/c
; Sequence 337, Application US/09818875
; Publication No. US20030051270A1
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
; FILE REFERENCE: Napro-4
; CURRENT APPLICATION NUMBER: US/09/818,875
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/192,176
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,179
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/208,538
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/244,989
```

```
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 4385
; SOFTWARE: Friedman macro Napro4
; SEQ ID NO 337
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-818-875-337
```

```
Query Match      18.7%; Score 23; DB 10; Length 121;
Best Local Similarity 54.0%; Pred. No. 6.4e+02;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCAGCTGACACAGGGCTAGACCCC 69
Db 118 CGTTAGTGTCTCCCTGGGGCAGCTCGTGTGAGGCTCCCCCTTTCTTTCGGGAGATTCTCT 59

QY 70 CTCACAGTGCGCCGCATGTGCTGCA 96
Db 58 TCCTCTGTGCGCCGGTCTCTCCACGGA 32
```

```
RESULT 4
US-09-818-875-338
; Sequence 338, Application US/09818875
; Publication No. US20030051270A1
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
; FILE REFERENCE: Napro-4
; CURRENT APPLICATION NUMBER: US/09/818,875
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/192,176
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,179
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/208,538
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/244,989
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 4385
; SOFTWARE: Friedman macro Napro4
; SEQ ID NO 338
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-818-875-338
```

```
Query Match      18.7%; Score 23; DB 10; Length 121;
Best Local Similarity 54.0%; Pred. No. 6.4e+02;
Matches 47; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 10 CGTTTCCTTCAGACTGAGGACAGCTGGGTCCAGCCTCAGCTGACACAGGGCTAGACCCC 69
Db 4 CGTTAGTGTCTCCCTGGGGCAGCTCGTGTGAGGCTCCCCCTTTCTTTCGGGAGATTCTCT 63

QY 70 CTCACAGTGCGCCGCATGTGCTGCA 96
Db 64 TCCTCTGTGCGCCGGTCTCTCCACGGA 90
```

```
RESULT 5
US-10-209-787-337/c
; Sequence 337, Application US/10209787
; Publication No. US20030217377A1
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 19:01:47 ; Search time 2458 Seconds
(without alignments)
4503.816 Million cell updates/sec

Title: US-10-057-813-13_COPY_1_273
Perfect score: 273
Sequence: 1 atggagcactaccggaagc.....ccaagctacgtttccttcag 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 55031126

Minimum DB seq length: 0
Maximum DB seq length: 273

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:*
118: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
119: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:*
120: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:*
121: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:*
122: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:*
123: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:*
124: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:*
125: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	226	82.8	264	13	US-08-856-624-2456
2	226	82.8	264	24	US-09-540-766-8863
3	218	79.9	271	14	US-08-941-869A-2841
4	218	79.9	271	24	US-09-540-208-35269
5	212	77.7	245	9	US-08-438-571A-2389
6	212	77.7	245	24	US-09-540-233D-126181
7	206	75.5	260	14	US-08-959-395-768
8	206	75.5	260	23	US-09-539-334-6888
9	196	71.8	262	13	US-08-812-505-1415
10	196	71.8	262	14	US-08-978-620-1415
11	196	71.8	262	24	US-09-540-229-45100
12	196	71.8	262	66	US-60-013-529-1415
13	193	70.7	250	14	US-08-951-197-5067
14	193	70.7	250	23	US-09-539-806-30498
15	193	70.7	250	23	US-09-539-806B-30498
16	193	70.7	250	67	US-60-027-249-1037
17	192	70.3	250	23	US-09-539-800-16108
18	192	70.3	250	23	US-09-539-800B-16108
19	192	70.3	250	23	US-09-539-800C-16108
20	190	69.6	244	15	US-09-014-441-1305
21	190	69.6	244	24	US-09-540-212A-30093
22	190	69.6	244	68	US-60-034-841-1305
23	189.6	69.5	264	24	US-09-540-213-18789
24	185	67.8	239	14	US-08-901-902-73
25	185	67.8	239	23	US-09-539-806-30137
26	185	67.8	239	23	US-09-539-806B-30137
27	185	67.8	239	67	US-60-023-379-73
28	184	67.4	233	13	US-08-879-863-3710
29	184	67.4	233	13	US-08-879-863A-3710
30	184	67.4	233	24	US-09-540-213-12679
31	183.4	67.2	259	18	US-09-293-657-1609
32	183.4	67.2	259	24	US-09-540-233D-76641
33	183.4	67.2	259	73	US-60-082-035-1609
34	183	67.0	238	23	US-09-539-334-28833
35	181	66.3	238	24	US-09-540-229-113556
36	169	61.9	227	9	US-08-413-150-22
37	169	61.9	227	24	US-09-540-499-10618
38	169	61.9	227	24	US-09-540-499B-10618
39	166	60.8	224	23	US-09-539-800-15324
40	166	60.8	224	23	US-09-539-800B-15324
41	166	60.8	224	23	US-09-539-800C-15324
42	166	60.8	234	14	US-08-923-902-2187
43	166	60.8	234	24	US-09-540-208-7903
44	157	57.5	210	13	US-08-801-504-2052
45	157	57.5	210	24	US-09-540-499-16308

ALIGNMENTS

RESULT 1
US-08-856-624-2456
; Sequence 2456, Application US/08856624
; GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
NUMBER OF SEQUENCES: 3070
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,624
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017,190
FILING DATE: MAY 13, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,100
FILING DATE: SEPTEMBER 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0174 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 2456:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 1929519
US-08-856-624-2456

Query Match 82.8%; Score 226; DB 13; Length 264;
Best Local Similarity 100.0%; Pred. No. 6.4e-51;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CCAATGCCCCAGCTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAAT 107
Db 1 CCAATGCCCCAGCTACCTCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAAT 60
QY 108 TCGCAACCTGCTGGGGTTGGGTCTGGGTCTGGGTGGAGGGCGGCGAGTCTCGGCAATGTAGT 167
Db 61 TCGCAACCTGCTGGGGTTGGGTCTGGGTCTGGGTGGAGGGCGGCGAGTCTCGGCAATGTAGT 120
QY 168 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGCAGTCTGGGTGAGATTGTCAAGCG 227
Db 121 GTTCTCAGGTTCTGGCAGGGCTGCAGGAAGGCTGCAGTCTGGGTGAGATTGTCAAGCG 180
QY 228 GCGGGTCCCAGGCCTGCACCGAGCTCACCAAGCTACGTTTCCTTCAG 273
Db 181 GCGGGTCCCAGGCCTGCACCGAGCTCACCAAGCTACGTTTCCTTCAG 225

US-09-540-766-8863
; Sequence 8863, Application US/09540766
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
; FILE REFERENCE: PD-1024 CIP
; CURRENT APPLICATION NUMBER: US/09/540,766
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
; NUMBER OF SEQ ID NOS: 77960
; SOFTWARE: PERL Program
; SEQ ID NO 8863
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00958497
US-09-540-766-8863

Query Match 82.8%; Score 226; DB 24; Length 264;
Best Local Similarity 100.0%; Pred. No. 6.4e-51;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CCCAATGCCCCAGCTACTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAAT 107
|||||
Db 1 CCCAATGCCCCAGCTACTCTGATACCCCTTGAGATGCGGGTCCGAGATGGCAGCAAAAT 60
|||||

QY 108 TCGCAACCTGCTGGGTTGGTCTGGGTGCGTTGGAGGCGGCGAGTGTCTCGGCATGTAGT 167
|||||
Db 61 TCGCAACCTGCTGGGTTGGTCTGGGTGCGTTGGAGGCGGCGAGTGTCTCGGCATGTAGT 120
|||||

QY 168 GTTCTCAGGTTCTGGCAGGCTGCAGGAAAGGCTGTACGCTGCGTGAGATTGTCAAGCG 227
|||||
Db 121 GTTCTCAGGTTCTGGCAGGCTGCAGGAAAGGCTGTACGCTGCGTGAGATTGTCAAGCG 180
|||||

QY 228 GCGGGTCCCAGGCTGCACAGCTCACCAAGCTACGTTTCCTTCAG 273
|||||
Db 181 GCGGGTCCCAGGCTGCACAGCTCACCAAGCTACGTTTCCTTCAG 226
|||||

RESULT 3
US-08-941-869A-2841
; Sequence 2841, Application US/08941869A
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN PROSTATE
; NUMBER OF SEQUENCES: 5486
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

US-08-941-869A-2841
; Sequence 2841, Application US/08/941,869A
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,839
; FILING DATE: SEPTEMBER 30, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/012,689
; FILING DATE: FEBRUARY 27, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0119-1 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 2841:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1794787
US-08-941-869A-2841

Query Match 79.9%; Score 218; DB 14; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 60
|||||
Db 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 113
|||||

QY 61 CTACCTCTGTATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
|||||
Db 114 CTACCTCTGTATACCTTGAGATGCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
|||||

QY 121 GGGTTGGCTCTGGGTGCGTTGGAGGCGGCGAGTGTCTGCGCATGTAGTGTCTCAGGTTCT 180
|||||
Db 174 GGGTTGGCTCTGGGTGCGTTGGAGGCGGCGAGTGTCTGCGCATGTAGTGTCTCAGGTTCT 233
|||||

QY 181 GGCAGGGCTGCAGGAAAGGCTGTCTGAGTGTCTGAGAT 218
|||||
Db 234 GGCAGGGCTGCAGGAAAGGCTGTCTGAGTGTCTGAGAT 271
|||||

RESULT 4
US-09-540-208-35269
; Sequence 35269, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 35269
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:


```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00317250
US-09-540-208-35269

Query Match      79.9%; Score 218; DB 24; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 60
Db 54 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 113
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 114 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 173
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGGCGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 174 GGGTTGGCTCTGGGTCGGTTGGAGGGGCGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 233
QY 181 GGCAGGGCTGCAGGAAAGGCTGTACGCTGCGCTGAGAT 218
Db 234 GGCAGGGCTGCAGGAAAGGCTGTACGCTGCGCTGAGAT 271
```

```
RESULT 5
US-08-438-571A-2389
; Sequence 2389, Application US/08438571A
; GENERAL INFORMATION:
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Levine, Wendy B.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Freitas, Nicole
; APPLICANT: Jernigan-Kelleher, Colleen
; APPLICANT: Stuart, Susan G.
; APPLICANT: Scott, Randal W.
; APPLICANT: Bills, Pamela Kay
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Akerblom, Ingrid Erika
; TITLE OF INVENTION: POLYNUCLEOTIDES DERIVED FROM THP-1 CELLS
; NUMBER OF SEQUENCES: 5094
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,571A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph.D.
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-001-5 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2389:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: P-033488
US-08-438-571A-2389

Query Match      77.7%; Score 212; DB 9; Length 245;
Best Local Similarity 98.1%; Pred. No. 4.1e-47;
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 60
Db 25 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCTTCCCCAATGCCCCAG 84
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db 85 CTACCTCCTGATACCCCTTNANATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 144
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGGCGCAGTNCCTCGGCATGTAGTGTCTCAGGTTCT 180
Db 145 GGGTTGGCTCTGGGTCGGTTGGAGGGGCGCAGTNCCTCGGCATGTAGTGTCTCAGGTTCT 204
QY 181 GGCAGGGCTGCAGGAAAGGCTGTACGCTGCGCTGAG 216
Db 205 GGCAGGGCTGCAGGAAAGGCTGTACGNTGCGCTGAG 240
```

```
RESULT 6
US-09-540-233D-126181
; Sequence 126181, Application US/09540233D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF HEMIC AND IMMUNE SYSTEM TISSUE
; FILE REFERENCE: PD-1030 CIP
; CURRENT APPLICATION NUMBER: US/09/540,233D
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 8/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/435,761
; PRIOR FILING DATE: 1995-05-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 129340
; SOFTWARE: PERL Program
; SEQ ID NO 126181
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00448643
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 103, 105, 178, 231
```

```
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-233D-126181

Query Match      77.7%; Score 212; DB 24; Length 245;
Best Local Similarity 98.1%; Pred. No. 4.1e-47;
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 60
   |||||||
Db 25 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 84

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
   |||||||
Db 85 CTACCTCCTGATACCCCTTNANATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 144

QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 216
   |||||||
Db 205 GGCAGGGCTGCAGGAAAGCTGTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 240

RESULT 7
US-08-959-395-768
; Sequence 768, Application US/08959395
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN PANCREATIC ISLET CELLS
; NUMBER OF SEQUENCES: 5789
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,395
; FILING DATE: HEREWITH
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,755
; FILING DATE: OCTOBER 28, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,551
; FILING DATE: DECEMBER 20, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0259 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 768:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 260 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 2071192H2
US-08-959-395-768

Query Match      75.5%; Score 206; DB 14; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 60
   |||||||
Db 49 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 108

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
   |||||||
Db 109 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 168

QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 180
   |||||||
Db 169 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 228

QY 181 GGCAGGGCTGCAGGAAAGGCTGTGTCAG 206
   |||||||
Db 229 GGCAGGGCTGCAGGAAAGGCTGTGTCAG 254

RESULT 8
US-09-539-334-6888
; Sequence 6888, Application US/09539334
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF ENDOCRINE SYSTEM TISSUE
; FILE REFERENCE: PD-1026 CIP
; CURRENT APPLICATION NUMBER: US/09/539,334
; CURRENT FILING DATE: 2000-03-30
; "Prior application data removed - refer to PALM or file wrapper"
; NUMBER OF SEQ ID NOS: 38381
; SOFTWARE: PERL Program
; SEQ ID NO 6888
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00887627
US-09-539-334-6888

Query Match      75.5%; Score 206; DB 23; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 60
   |||||||
Db 49 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCCCAATGCCCCCAG 108

QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
   |||||||
Db 109 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 168

QY 121 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 180
   |||||||
Db 169 GGGTTGGCTCTGGTCTGGTGGAGGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 228

QY 181 GGCAGGGCTGCAGGAAAGGCTGTGTCAG 206
   |||||||
Db 229 GGCAGGGCTGCAGGAAAGGCTGTGTCAG 254
```

RESULT 9
US-08-812-505-1415
; Sequence 1415, Application US/08812505
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: ASTROCYTOMA
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,505
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/013,529
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-812-505-1415

Query Match 71.8%; Score 196; DB 13; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCTCCCGCCCTTCCCCAATGCCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGCTCCCGAGCTCCCGCCCTTCCCCAATGCCCCCAG 119
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
120 CTACCTCCTGATACCCCTTNANATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGGTCCGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
180 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGGTCCGGCATGTAGTGTCTCAGGTTCT 239
QY 181 GGCAGGGCTGCAGGAAGGCT 201
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
240 GGCAGGGCTGCAGGAAGGCT 260

RESULT 10
US-08-978-620-1415
; Sequence 1415, Application US/08978620
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN BRAIN
; NUMBER OF SEQUENCES: 7797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,620
; FILING DATE: HERewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,151
; FILING DATE: DECEMBER 6, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/812,505
; FILING DATE: MARCH 7, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
US-08-978-620-1415

Query Match 71.8%; Score 196; DB 14; Length 262;
Best Local Similarity 97.5%; Pred. No. 9.5e-43;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCCGAGCTCCCGCCCTTCCCCAATGCCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
60 ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGCTCCCGAGCTCCCGCCCTTCCCCAATGCCCCCAG 119
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
120 CTACCTCCTGATACCCCTTNANATGCGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 179
QY 121 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGGTCCGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
180 GGGTTGGCTCTGGTTCGGTTGGAGGGCGGAGTGGTCCGGCATGTAGTGTCTCAGGTTCT 239

Qy	181	GGCAGGGCTGCAGGAAGGCT	201
Db	240	GGCAGGGCTGCAGGAAGGCT	260

```

RESULT 11
US-09-540-229-45100
; Sequence 45100, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 45100
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00691274
; NAME/KEY: unsure
; LOCATION: 38, 76, 87, 138, 140, 168
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-45100

```

Query Match	71.8%	Score 196;	DB 24;	Length 262;
Best Local Similarity	97.5%	Pred. No. 9.5e-43;		
Matches 196;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY 1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 60			
Db 60	ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGCTCCAGCGCCTTCCCAATGCCCCAG 119			
QY 61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCCGAACCTGCTG 120			
Db 120	CTACCTCCTGATACCCCTTNANATCGGGTCCGAGATGGCAGCAAAATTNGCAACCTGCTG 179			
QY 121	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 180			
Db 180	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGCTCGGCATGTAGTGTCTCAGGTTCT 239			
QY 181	GGCAGGGCTGCAGGAAAGGCT 201			
Db 240	GGCAGGGCTGCAGGAAAGGCT 260			

RESULT 12
US-60-013-529-1415
; Sequence 1415, Application US/60013529
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: ASTROCYTOMA
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

```

; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/013,529
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 1415:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; CLONE: 862319
;
; US-60-013-529-1415

```

	Query Match	71.8%;	Score 196;	DB 66;	Length 262;
	Best Local Similarity	97.5%;	Pred. No. 9.5e-43;		
	Matches 196;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1	ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG	60		
DB	60	ATGGAGCACTACCGGANAGCTGGCTCTNTAGAGTCCCAGCGCCTTCCCCAATGCCCCAG	119		
QY	61	CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTGCAACCTGCTG	120		
DB	120	CTACCTCCTGATACCCCTTNANATCGGGTCCGAGATGGCAGCAAAATTNGCAACCTGCTG	179		
QY	121	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTTCTCAGGTTCT	180		
DB	180	GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCACTGCTCGGCATGTAGTGTTCTCAGGTTCT	239		
QY	181	GGCAGGGCTGCAGGAAAGGCT	201		
DB	240	GGCAGGGCTGCAGGAAAGGCT	260		

RESULT 13
US-08-951-197-5067
; Sequence 5067, Application US/08951197
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuv , Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM HUMAN
; TITLE OF INVENTION: BREAST
; NUMBER OF SEQUENCES: 5970
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951.197
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,249
; FILING DATE: OCTOBER 1, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/826,438
; FILING DATE: MARCH 20, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,285
; FILING DATE: MARCH 20, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0143-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5067:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1997229
; US-08-951-197-5067

Query Match 70.7%; Score 193; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 117
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGGCATGTAGTGTCTCAGGTTCT 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGGCATGTAGTGTCTCAGGTTCT 237
QY 181 GGCAGGGCTGCAG 193
Db |||||||||||||||
238 GGCAGGGCTGCAG 250

RESULT 14
US-09-539-806-30498
; Sequence 30498, Application US/09539806
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE

; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539,806
; CURRENT FILING DATE: 2000-03-30
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 48372
; SOFTWARE: PERL Program
; SEQ ID NO 30498
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00555791
; US-09-539-806-30498

Query Match 70.7%; Score 193; DB 23; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGTCCCGAGCGCCTTCCCCAATGCCCCAG 117
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177
QY 121 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGGCATGTAGTGTCTCAGGTTCT 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178 GGGTTGGCTCTGGTTCGGTGGAGGGCGGCGAGTGGCATGTAGTGTCTCAGGTTCT 237
QY 181 GGCAGGGCTGCAG 193
Db |||||||||||||||
238 GGCAGGGCTGCAG 250

RESULT 15
US-09-539-806B-30498
; Sequence 30498, Application US/09539806B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539,806B
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/706,766
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/004,676
; PRIOR FILING DATE: October 2, 1995
; PRIOR APPLICATION NUMBER: 08/749,515
; PRIOR FILING DATE: November 15, 1996
; PRIOR APPLICATION NUMBER: 60/006,810
; PRIOR FILING DATE: November 15, 1995
; PRIOR APPLICATION NUMBER: 08/822,285
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/013,696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/951,197
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027,249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/826,438
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/016,145
; PRIOR FILING DATE: April 18, 1996
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 48372
; SOFTWARE: PERL Program

; SEQ ID NO 30498
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00555791
US-09-539-806B-30498

Query Match 70.7%; Score 193; DB 23; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.2e-42;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCCTTCCCCAATGCCCCAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
58 ATGGAGCACTACCGGAAAGCTGGCTCTGTAGAGCTCCAGCGCCCTTCCCCAATGCCCCAG 117
QY 61 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 CTACCTCCTGATACCCCTTGAGATCGGGTCCGAGATGGCAGCAAAATTCGCAACCTGCTG 177
QY 121 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
178 GGGTTGGCTCTGGGTCGGTTGGAGGGCGGCAGTGTCTGGCATGTAGTGTCTCAGGTTCT 237
QY 181 GGCAGGGCTGCAG 193
Db ||||||||||||
238 GGCAGGGCTGCAG 250

Search completed: January 31, 2005, 20:58:10
Job time : 2461 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 21:02:13 ; Search time 68 Seconds
(without alignments)
1285.692 Million cell updates/sec

Title: US-10-057-813-13_COPY_253_375
Perfect score: 123
Sequence: 1 accaagctacgttctctca.....tgctgctcagccgggaacccc 123

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1041130

Minimum DB seq length: 0
Maximum DB seq length: 123

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		SUMMARIES			
Result No.	Score	% Query Match		DB ID	Description
		Match	Length		
C 1	23	18.7	89	4	US-09-513-999C-15513 Sequence 15513, A
C 2	21.4	17.4	119	4	US-09-513-999C-15214 Sequence 15214, A
C 3	20.6	16.7	77	4	US-09-513-999C-24985 Sequence 24985, A
C 4	20.4	16.6	66	4	US-09-513-999C-15004 Sequence 15004, A
C 5	20.4	16.6	90	4	US-08-120-324-13 Sequence 13, Appli
C 6	20.2	16.4	106	4	US-09-513-999C-33439 Sequence 33439, A
C 7	20.2	16.4	116	4	US-09-513-999C-13916 Sequence 13916, A
C 8	19.8	16.1	112	4	US-09-513-999C-14651 Sequence 14651, A
C 9	19.6	15.9	92	4	US-09-513-999C-25968 Sequence 25968, A
C 10	19.6	15.9	103	4	US-09-513-999C-15287 Sequence 15287, A
C 11	19.6	15.9	103	4	US-09-513-999C-15605 Sequence 15605, A
C 12	19.4	15.8	78	3	US-09-463-380-7 Sequence 7, Appli
C 13	19.4	15.8	78	4	US-09-985-357A-7 Sequence 7, Appli
C 14	19.2	15.6	77	2	US-08-477-527A-227 Sequence 227, App
C 15	19.2	15.6	77	3	US-08-481-710-227 Sequence 227, App
C 16	19.2	15.6	77	5	PCT-US96-09537-227 Sequence 227, App
C 17	19.2	15.6	91	4	US-09-270-767-27356 Sequence 27356, A
C 18	19	15.4	88	4	US-09-322-357-48 Sequence 48, Appli
C 19	19	15.4	89	4	US-09-621-976-14962 Sequence 14962, A
C 20	19	15.4	112	4	US-09-702-705-1575 Sequence 1575, Ap
C 21	19	15.4	112	4	US-09-736-457-1575 Sequence 1575, Ap
C 22	19	15.4	112	4	US-09-614-124B-1575 Sequence 1575, Ap
C 23	19	15.4	112	4	US-09-671-325-1575 Sequence 1575, Ap
C 24	19	15.4	112	4	US-09-658-824-1575 Sequence 1575, Ap
C 25	18.8	15.3	96	6	5185259-5 Patent No. 5185259
C 26	18.6	15.1	60	3	US-08-828-712-2 Sequence 2, Appli
C 27	18.6	15.1	60	3	US-09-063-276-2 Sequence 2, Appli

C 28	18.6	15.1	60	4	US-09-324-782-2	Sequence 2, Appli
C 29	18.6	15.1	60	4	US-09-668-143-2	Sequence 2, Appli
C 30	18.6	15.1	90	3	US-08-952-793-176	Sequence 176, App
C 31	18.6	15.1	90	4	US-09-849-928-176	Sequence 176, App
C 32	18.6	15.1	90	5	PCT-US96-09455A-176	Sequence 176, App
C 33	18.6	15.1	95	4	US-09-513-999C-20555	Sequence 20555, A
C 34	18.6	15.1	99	4	US-09-513-999C-32775	Sequence 32775, A
C 35	18.6	15.1	100	3	US-09-298-886-20	Sequence 20, Appl
C 36	18.6	15.1	100	4	US-09-999-672-20	Sequence 20, Appl
C 37	18.4	15.0	55	4	US-08-956-171E-2677	Sequence 2677, Ap
C 38	18.4	15.0	55	4	US-08-781-986A-2677	Sequence 2677, Ap
C 39	18.4	15.0	96	1	US-08-190-802A-20	Sequence 20, Appl
C 40	18.4	15.0	96	3	US-08-477-346-20	Sequence 20, Appl
C 41	18.4	15.0	96	3	US-08-473-089-20	Sequence 20, Appl
C 42	18.4	15.0	96	4	US-08-487-072A-20	Sequence 20, Appl
C 43	18.4	15.0	107	4	US-09-513-999C-35019	Sequence 35019, A
C 44	18.4	15.0	111	4	US-09-513-999C-12508	Sequence 12508, A
C 45	18.4	15.0	122	4	US-09-513-999C-27085	Sequence 27085, A

ALIGNMENTS

RESULT 1
US-09-513-999C-15513/c
; Sequence 15513, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15513
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15513

Query Match	18.7%	Score 23;	DB 4;	Length 89;
Best Local Similarity	60.3%	Pred. No. 1.8e+02;		
Matches	38;	Conservative	0;	Mismatches 25; Indels 0; Gaps 0;
QY	60	GCTAGACCCCTCACAGTGGCGGCCCATGTGCCTGCAGTGTGGGTGCTCAGCCGGGA	119	
Db	70	GCTCGTCCGTCTCCGCTGCGCCAGCCGCCGCTACTGCTGCTGCTCTCTCCGGAT	11	
QY	120	CCC 122		
Db	10	GCC 8		

RESULT 2
US-09-513-999C-15214/c
; Sequence 15214, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15214
; LENGTH: 119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15214

Query Match 17.4%; Score 21.4; DB 4; Length 119;
Best Local Similarity 61.8%; Pred. No. 6.1e+02;
Matches 34; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 69 CCTCACAGTGGCGCCCATGTGCTGCAGTGTGGTGTGCTGCTCAGCCGGGACCCC 123
Db 76 CATGAACGTGGGCAGGAAGCGGCTGCAATGCTGGGGCTGTCTGTCATGGAGCGC 22

RESULT 3
US-09-513-999C-24985
; Sequence 24985, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24985
; LENGTH: 77
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-24985

Query Match 16.7%; Score 20.6; DB 4; Length 77;
Best Local Similarity 74.3%; Pred. No. 1e+03;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 26 AGGACAGTGGTCCCGCCTCACCTGACACAGGG 60
Db 32 AGGACTTCTGTTTCAGCATAACCTAAGACAGGG 66

RESULT 4
US-09-513-999C-15004/c
; Sequence 15004, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15004
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15004

Query Match 16.6%; Score 20.4; DB 4; Length 66;
Best Local Similarity 58.1%; Pred. No. 1.1e+03;
Matches 36; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 35 GGGTCCCGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGCTG 94
Db 65 GGTGCGCTGCATTCTTTTCACACTGTCCACAGCCTCACATACCGGCTCCGTGTGGCTA 6

QY 95 CA 96
Db 5 CA 4

RESULT 5
US-08-120-324-13
; Sequence 13, Application US/08120324
; Patent No. 6723556
; GENERAL INFORMATION:
; APPLICANT: Saxena, Brij B.
; APPLICANT: Rathnam, Premila
; TITLE OF INVENTION: hCG-hLH RECEPTOR AND hCG-hLH
; TITLE OF INVENTION: RECEPTOR-hCG COMPLEX AS ANTIGENS, ANTIBODIES THERETO AND
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/120,324
; FILING DATE: 14-SEP-1993
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,613
; FILING DATE: 11-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/879,245
; FILING DATE: 06-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/742,236
; FILING DATE: 08-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/555,696
; FILING DATE: 23-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/910,554
; FILING DATE: 23-SEP-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/752,497
; FILING DATE: 08-JUL-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/446,145
; FILING DATE: 02-DEC-1982
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: A6229-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-120-324-13

Query Match      16.6%; Score 20.4; DB 4; Length 90;
Best Local Similarity 61.1%; Pred. No. 1.2e+03;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 67 CCCCTCACAGTGGCGCGCATGTGCCCTGCAGTGTGGTGTGCTGCTCAGCCGGGAC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CTCCGCCAGGGGGCGCGACGTGCGAGCGCTGAGCGAGCTGCAAGGGCGCGCC 54

RESULT 6
US-09-513-999C-33439
; Sequence 33439, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 33439
; LENGTH: 106
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-33439

Query Match      16.4%; Score 20.2; DB 4; Length 106;
Best Local Similarity 63.3%; Pred. No. 1.4e+03;
Matches 31; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 72 CACAGTGGCGCGCATGTGCCTGCAGTGTGGTGTGCTGCTCAGCCGGGAC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 46 CCCGGCGGGGGCCTGGTGACGGGAGTGTGCTGTCTCTGGAC 94

RESULT 7
US-09-513-999C-13916
; Sequence 13916, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13916
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-13916

Query Match      16.4%; Score 20.2; DB 4; Length 116;
Best Local Similarity 54.8%; Pred. No. 1.5e+03;
Matches 40; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 49 CCTGACACAGGGCTAGACCCCTCACAGTGGCGCGCATGTGCTGCAGTGTGGGTGCTG 108
```

```

   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44 CCGGAGATGGCGCGGTGCTGAACGCAGACGAGCTCGAGGTGTCCGTGACGGCCTCACG 103

QY 109 CTCAGCCGGGACC 121
   ||||| ||||| |||||
Db 104 CTCAGCCCGGACC 116

RESULT 8
US-09-513-999C-14651
; Sequence 14651, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14651
; LENGTH: 112
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 65
; OTHER INFORMATION: r=a or g
US-09-513-999C-14651

Query Match      16.1%; Score 19.8; DB 4; Length 112;
Best Local Similarity 55.4%; Pred. No. 1.9e+03;
Matches 36; Conservative 1; Mismatches 28; Indels 0; Gaps 0;

QY 20 AGACTGAGGACAGCTGGTCCCGAGCTCACCTGACACAGGGCTAGACCCCTCACAGTGC 79
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 AGACTTCAGGAGACCARAGCCCGAGCTTGCCAGGCACTGAGCTAGAGCCCTGCCATGGC 107

QY 80 GCCGC 84
   |||
Db 108 ACCCC 112

RESULT 9
US-09-513-999C-25968
; Sequence 25968, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 25968
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6
; OTHER INFORMATION: h=a or c or t
```

```
;
; NAME/KEY: misc_feature
; LOCATION: 7
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: v=a or c or g
US-09-513-999C-25968
```

```
Query Match          15.9%; Score 19.6; DB 4; Length 92;
Best Local Similarity 60.4%; Pred. No. 2.2e+03;
Matches 29; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 75 AGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTGCTCAGCGGGACCC 122
      ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATTGCHNNVTGATGTCCCTGCTGTGGTATCCTGGCAGGTCCCAACCC 48
```

```
RESULT 10
US-09-513-999C-15287/c
; Sequence 15287, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15287
; LENGTH: 103
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3
; OTHER INFORMATION: k=g or t
US-09-513-999C-15287
```

```
Query Match          15.9%; Score 19.6; DB 4; Length 103;
Best Local Similarity 58.6%; Pred. No. 2.2e+03;
Matches 34; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 66 CCCCCTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTCAGCGGGACCC 123
      ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 CCGCCTGCCTTGGCCTCCCAAGTGTGAGATTGCAGCCTCTGCCAGCGCCACCC 16
```

```
RESULT 11
US-09-513-999C-15605/c
; Sequence 15605, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
```

```
;
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15605
; LENGTH: 103
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: v=a or c or g
US-09-513-999C-15605
```

```
Query Match          15.9%; Score 19.6; DB 4; Length 103;
Best Local Similarity 58.6%; Pred. No. 2.2e+03;
Matches 34; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 66 CCCCCTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTCAGCGGGACCC 123
      ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 CCGCCTGCCTTGGCCTCCCAAGTGTGAGATTGCAGCCTCTGCCAGCGCCACCC 16
```

```
RESULT 12
US-09-463-380-7/c
; Sequence 7, Application US/09463380
; Patent No. 6391633
; GENERAL INFORMATION:
; APPLICANT: Anne STERN; Michael BRANDT; Konrad HONOLD; Johannes AUER; Hans KOLL
; TITLE OF INVENTION: Preparation of erythropoietin by endogenous gene activation
; FILE REFERENCE: HUBER 1151 PFF/MAS
; CURRENT APPLICATION NUMBER: US/09/463,380
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/EP98/04590
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: US 09/113,692
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: DE 19753681.1
; PRIOR FILING DATE: 1997-12-03
; PRIOR APPLICATION NUMBER: EP 97112640
; PRIOR FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Wordperfect 6/7/8
; SEQ ID NO 7
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..60
; OTHER INFORMATION: Nucleotide sequence of the primer EX3 (Met-Gly-Ala-His).
US-09-463-380-7
```

```
Query Match          15.8%; Score 19.4; DB 3; Length 78;
Best Local Similarity 64.4%; Pred. No. 2.4e+03;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 70 CTCACAGTGGCGCGCCATGTGCCTGCAGTGTGGGTGCTGCTCAGC 114
      ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 CTCACCGTGGGCGCCCATCTCCGCGCCTTGGCGGGGTCCCTCAGC 22
```

```
RESULT 13
US-09-985-357A-7/c
; Sequence 7, Application US/09985357A
; Patent No. 6544748
; GENERAL INFORMATION:
; APPLICANT: Anne STERN; Michael BRANDT; Konrad HONOLD; Johannes AUER; Hans KOLL
; TITLE OF INVENTION: Preparation of erythropoietin by endogenous gene activation
; FILE REFERENCE: HUBER 1151.1 CON PFF/MAS
; CURRENT APPLICATION NUMBER: US/09/985,357A
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 09/463,380
; PRIOR FILING DATE: 2000-01-21
```

;; PRIOR APPLICATION NUMBER: PCT/EP98/04590
;; PRIOR FILING DATE: 1998-07-22
;; PRIOR APPLICATION NUMBER: US 09/113,692
;; PRIOR FILING DATE: 1998-07-10
;; PRIOR APPLICATION NUMBER: DE 19753681.1
;; PRIOR FILING DATE: 1997-12-03
;; PRIOR APPLICATION NUMBER: EP 97112640
;; PRIOR FILING DATE: 1997-07-23
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: Wordperfect
;; SEQ ID NO 7
;; LENGTH: 78
;; TYPE: DNA
;; ORGANISM: Artificial sequence
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 49..60
;; OTHER INFORMATION: Nucleotide sequence of the primer EX3 (Met-Gly-Ala-His).
US-09-985-357A-7

Query Match 15.8%; Score 19.4; DB 4; Length 78;
Best Local Similarity 64.4%; Pred. No. 2.4e+03;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 70 CTCACAGTGGCGCCGATGTGCCTGCAGTGTGGTGCTGCTCAGC 114
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 CTCACGTGGCGCCCATCTCCGCGCCTGGCGGGGTCCCTCAGC 22

RESULT 14
US-08-477-527A-227/c
; Sequence 227, Application US/08477527A
; Patent No. 5972599
; GENERAL INFORMATION:
; APPLICANT: DIANE TASSET
; APPLICANT: NIKOS PAGRATIS
; APPLICANT: SUMEDHA JAYASENA
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: OF CYTOKINES
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,527A
; FILING DATE: 7-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/931,473
; FILING DATE: 17-AUGUST-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: 8-SEPTEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane H. McClearn

;; REGISTRATION NUMBER: 33,960
;; REFERENCE/DOCKET NUMBER: NEX41-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (303) 793-3333
;; TELEFAX: (303) 793-3433
;; INFORMATION FOR SEQ ID NO: 227:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 77 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-477-527A-227

Query Match 15.6%; Score 19.2; DB 2; Length 77;
Best Local Similarity 54.2%; Pred. No. 2.8e+03;
Matches 39; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 32 GCTGGGTCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGC 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 GCCGATCCGGCCCTCATGTGCAACAGTGCAGCTGAGCGGACTGCGCATCGACTTGAGC 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 92 CTGCAGTGTGGG 103
Db 17 GTTTATTCTGAG 6

RESULT 15
US-08-481-710-227/c
; Sequence 227, Application US/08481710
; Patent No. 6028186
; GENERAL INFORMATION:
; APPLICANT: DIANE TASSET
; APPLICANT: NIKOS PAGRATIS
; APPLICANT: SUMEDHA JAYASENA
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: OF CYTOKINES
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,710
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/931,473
; FILING DATE: 17-AUGUST-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: 8-SEPTEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane H. McClearn
; REGISTRATION NUMBER: 33,960


```
;
; REFERENCE/DOCKET NUMBER: NEX41-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-481-710-227
```

```
Query Match 15.6%; Score 19.2; DB 3; Length 77;
Best Local Similarity 54.2%; Pred. No. 2.8e+03;
Matches 39; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 32 GCTGGTCCAGCCTCACCTGACACAGGGCTAGACCCCTCACAGTGGCGCCCATGTGC 91
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 GCCGATCCGGGCTCATGTGGAACAGTGACGCTGAGGGGACTGCGCATCGACTTGAGC 18

QY 92 CTGCAGTGTGG 103
Db | | | | |
17 GTTATTCTGAG 6
```

Search completed: January 31, 2005, 22:17:25
Job time : 69 secs